UNIVERSITY OF THE PELOPONNESE
SCHOOL OF ECONOMICS, MANAGEMENT AND INFORMATICS
DEPARTMENT OF INFORMATICS AND TELECOMMUNICATIONS

Knowledge and Uncertainty Research Laboratory

Undergraduate Thesis

Automation of the research impact estimation process

Bampetas Nikolaos
2025201000043

Supervisor:

Wallace Manolis
Lecturer

Tripolis, June 2015

Approved by the comitee of Inquiry on January 31, 2015.

Giorgos Lepouras	Nikolaos Tselikas
Associate professor	Assistant professor

.....

Nikolaos Bampetas
Department of Information Technology and Telecommunications
University of Peloponnese

This work was in collaboration with the Research Group Knowledge and Uncertainty (ΓΑΒ LAB)

Abstract

It is more than a fact that humanity has the tendency to develop, thus driving scientists to conducting more research. So the number of researchers has increased considerably and so has the number of researches that are published. To add to that every paper has its own citations that refer to others already published researches that in turn reference previous ones, creating the need for a tool that will be able to collect all those lists from a citation database, accompanied with all the details like names of authors, publishers and citations.

In this thesis we create new software that collects a list of the citations of a researcher.

Our software draws information from Google Scholar, processes that information in order to identify self-citations and then exports in a generic and easily re-wable XML format.

Περίληψη

Στις μέρες μας, η ανθρωπότητα έχει εξελιχθεί και αυτο μας οδήγησε σε περισσότερες έρευνες. Έτσι ο αριθμός των ερευνητών αυξήθηκε αρκετά οπως και οι έρευνες τους. Επιπλέον κάθε αναφορά έχει τις δικές του αναφορές που επίσης χρειαζόμαστε τους τίτλους τους. Παρόλα αυτα ακόμα δεν υπάρχει μια εφαρμογή που να μπορεί να μαζέψει ολους αυτους τους τίτλους απο μια βάση αναφορών, με όλες τις πληροφορίες που θέλουμε οπως ερευνητές, εκδότες και αναφορές.

Σε αυτη την πτυχιακή σκοπός μας είναι να φτιάξουμε μια τέτοια εφαρμογή που θα συλλέγει όλες τις λίστες με τις αναφορές για καθένα ερευνητή.

Το πρόγραμμα αντλεί πληροφορίες απο το Google Scholar, επεξεργάζεται αυτη την πληροφορία με σκοπό να εντοπίσει τις προσωπικές αναφορές και τέλος τα προσφέρει σε μια εύκολη μορφή XML τυπου

Dedicated to those who helped me
in this project.

Contents

Preface

0.1 Acknowledgments

I am grateful to Manolis Wallace, lecturer and Gav Lab members. I am extremly thankful and indebted to them for sharing expertise, and sincere and valuable guidance and encouragement extended to me.

Chapter 1

Introduction

It is a fact that the 20th century is actually the time when computer science has evolved astoundingly in all fields. One of the results of such an evolution is that everything can be published on the Internet thus making anything easier to be called upon when the need arises. Scientific papers are written in connection to a certain research and are published in a journal or presented in a conference.These publications can also be found on the internet on some specific databases in order to assist in education and research procedure. This has proven to be very useful to the academic world, because everyone can be informed about the innovations in his scientific field, or build a strong background in a field of a certain science and in this way contribute in one's own evolution. But it is not only the papers that play a significant role in the academic world but also theirs citations and that is because they can become a measure of a researcher's impact on science.

Currently there are applications that can easily provide us with a list of the published papers of a researcher just by giving his full name as input. Most of these can give us results with not great accuracy and precision like JabRef retrieves only the first page that Google Scholar returns. Moreover, none of the existing applications have the ability to offer us all citations of a particular paper. To be more precise, all this information does exist and is offered free in databases that anyone can find on Internet, but it seems to be extremely difficult and requiring a lot of work hours to retrieve all this information by hand searching , which apparently is the only way to achieve it.

This thesis is investigating a proper way for someone to retrieve a full list of paper titles that are published as well as the year they were published and other details. Moreover it will retrieve the citations for each publication of a particular researcher and finally to export all this information in the form of an XML document. In the next chapters we will analyze the importance of an academic paper and its role in the academic society. We will also deal with the places that someone can find such information, the databases, and the tools it can be retrieved with. After that we will mention,the way our citation software was created, how we store our data, where it was retrieved from an by which method.This thesis provides the scope of this project , offering a concise view of, why we are going to create an application that retrieves papers and their citations and how this will help scientific society.

Specifically in chapter 2 provides information about the literature review, as well as what a scientific paper and citations are, where citations can be found meaning in which databases. In addition chapter two will give us information about the citations programs that already exist and how they collect papers and their citations.

In chapter 3 we discuss our contribution to this project but mostly the problems that he was faced with. After that, we will propose some methodology that will help in surpassing these prob-

lems and we will give a work plan on how we had worked to complete our project.

Chapter 4 provides information on how we have made the data retrieval from google scholar, which were the difficulties during our research.

Chapter 5 investigates the data processing , the data structures that are being used by the software and how we had managed to delete the self-citations.

Chapter **??** is shows our application how it's work and the user interface.

The seventh chapter projects the trials and the results of our final application.

Finally **??** summarizes our work and our future plans

Chapter 2

Literature Review

It is often said that academic society is responsible for the evolution of technology and the existence of modern, comfortable life. It is also mentioned that development in various fields of expertise is caused by researchers , theirs ideas and obviously by their work. These have to be made publicly known in order to spread the knowledge of research, and this is achieved by publishing researchers' work. Professors write a report that concludes their research and then publish it. The publication can be done in journals, in conferences and all that ends up on the Internet and in specific databases. As we can understand every single researcher contributes in scientific progress. There are also citations to each published paper. Citations can offer important information on the importance of the specific paper and this can be measured via citation metrics.

If we have all the citations of a researcher we can count the h-index , the g-index which are the best way to have a precise number about the impact of a researcher. Specifically h-index tries to compute the productivity and the citation impact of a researcher. The g -index i used for quantifying scientific productivity which is based on publication number. To conclude there are a lot of metrics in order to compute all the impact of the work of an individual researcher. In order to calculate all the above metrics, we have to not conclude the self-citation. Self citation is a published paper that has the same author with the first paper. This citations can not be included in our calculation because as we can understand,they does not really give real importance to the researcher. This is why in the next chapter we will show a method to separate citations and self-citations.

2.1 Papers and their Citations

Researchers use papers in order to publish their work concerning a particular project. After the publication other researchers that want to get informed on the evolution in this field of science can access via internet the paper itself ,some times for free and some others by paying . As we can understand papers play an important role in society because every evolution in any sector takes place after research has been done. This research is published in a paper. This paper is sure to contain citations as citations are the way to inform the reader that a part of a research is based on someone else's work.

Citations are quite important for academic purposes, and can be used in order to calculate some metrics and find out how important each research is. Citation Analysis in simple words is the method to link a research work with an other one and more specifically it examines the frequency of citations in articles, books and papers.

The examination of citations can show us when examining two papers which one is more

important. It can provide information on the quality of the research of a researcher and to determine how much impact a particular author has made by looking at his total number of citations. So citations are a way to calculate the impact of a certain researcher on science. This is that has what led us to the decision of creating such a software which will give out a list of all the citations of the paper.

2.2 Citation databases

The 20th century presented a huge increase of the number of researches, which is the basic reason for the creation of tools in order to save and retrieve at any time all papers and citations that we need. Moreover these tools should be user friendly and have the ability to collect a list of the work of any particular researcher by just typing his name . We understand that the results should be accurate and precise. These tools are supported by Internet databases that have already linked the citations with the paper they are mentioned in. Some of them are Google scholar, web of science, academic search, Cite Seer and Scopus.

Google scholar as we can figure out from its name, is a tool that Google offers free for the academic society. This database stores all the information that we need in our research. The data mining from Google scholar is quite easy to do but the most importantly it can also provide s a big variety of a researcher's work. Google scholar is one of the top citation databases and it is frequently used by users as in May of 2014 there were 160 million documents in it and it is continually updated. It can provide the users with all the information about a certain researcher like the title of the paper or book, the date of publication ,the publisher and the co-authors .In addition, the citations of each paper can be found next to it in a link that redirects us to a new list with all the papers that were written based on the previous one. Google scholar is mostly a search tool for academic papers and not exactly a database, but this is what we need in our research.

Scopus ,on the other hand, is not a free use database. Scopus is a service of Elsevier and is considered to be the biggest citation database with over 20.000 titles, 5000 publishers and the extreme number of 5 million papers previously presented in conferences.

Web of Science is a well known tools for citations and for papers and it is part of the platform web of knowledge. Web of Science provides fast and precise access to citation databases and might have less information than Scopus but it entails about 150.000 conference papers concerning three different scientific fields. Web of Science has the Science Citation index Expanded with details on 6.500 papers for supplied sciences since 1970 up until now, the Social Sciences citation index with 1.950 papers on social sciences and regarding art and human science it provides the Arts and human citation index on about to 1.000 papers.

Academic search was created by Microsoft in order to help young students and researchers to find the papers or the citations that are looking for easily. Academic search does not only retrieve papers but notices the strong points and the relations between papers.

Finally CiteSeer is also a citation database in addition to being an academic search engine like all the above databases . Cite Seer has in its collection documents of academic and science researches. It got widely known in 1998 and it creates a list of links that can be in some way a good means for the assessment of papers . CiteSeer works like Google scholar and it uses links that lead to the expected document. In addition to that all the metrics are computed for all the documents that are part of the lists which is returned by cite seer. A table with those citation databases and some informations about them

NAME	DESCRIPTION	OWNER
Google scholar	Multidisciplinary	Google
Web of science	Science (General)	Thomson Reuters
Scopus	Multidisciplinary	Elsevier
Academic search	Multidisciplinary	EBSCO Publishing
CiteSeer	Computer Science	Pennsylvania State University

2.3 Citation software

Citation software are an extension of citation databases, and that is because they mine the data of those databases in order to collect a full list sof all the titles of papers of a researcher. In addition some of them have the option of selecting the database the user wants to retrieve these lists from. These software available are Publish or Perish and JabRef. These two applications can provide precise lists of the published papers titles and moreover they provide the user with related information such as where those papers were published, which year and who the authors were.

Publish or Perish is a software that is powered by harzing.com. It is compatible with linux, macos and windows, and it is not an open source. It uses its own google scholar or microsoft Academic search as a database and gives the option of exporting the results of a search. It also displays information about the metrics of the researcher and the number of their work. This information are also readily provided by google scholar.

JabRef is an open source application. It can mine data from many citation dabases like diVa, Google scholar,IEEEXplore, INSIPIRE, medline and others. The results of jabref are not very accurate in comparison to publish or perish which can easily be spotted if someone uses both of them. JabRef can get only the first page of google scholar in contrast to publish or perish that returns the full list of the papers.

Citation software are very important but there are only two . In addition these two have the power to retrieve only the papers of a researcher and not their citations. In a previous section of this chapter we analyzed how important the citations are for a researcher and this is the basic reason of this project. We are going to create an application that will retrieves, firstly a list of the papers that a researcher published, and secondly a list for each paper with the relative citations.

Chapter 3

Contribution

To sum up, researchers print their thoughts, their work and their results in academic papers. Then anyone can expand on these researches on a new paper which will cite to the first one. If anyone wants to find a list of these documents he can find one in Google scholar with a good amount of precision. There are also some applications that provide you with all this information except for the citations. So our contribution is that we are going to make a program that will retrieve the published papers of a researcher as well as their citations. Finally we will have the option of exporting all this data in an XML scheme.

3.1 Problem description

As we saw in previous sections of this thesis papers and citations are very important for the academic society for a lot of reasons, and there are quite a few databases where one can find them. The basic problem is that people need a precise and accurate list of the published papers of a researcher and their citation as well. Until now there is not any particular software that can provide us with such a list. On the other hand there are some databases as Google scholar which gives as a precise list of all the above but the retrieval has to be done by hand. This procedure is very time consuming so we have to solve this problem by creating a new software. But the most important problem is that there is a big amount of citation in some publications and their retrieval can take months.

3.2 Proposed methodology

The proposed methodology to erase this problem is to create an application that will retrieve and export all the information that we need. This will become feasible by creating a program that will use Google scholar as a database. It will give google scholar the name of the researcher. Google scholar will then give us an html page that contains a full list of the published papers of the researcher that we gave as input. After that our application will parse this html page and will separate the html code from the tags that contain the needed information. This will then lead to analyzing the acquired information, deleting special characters and spaces and saving this information. Finally it should provide to the user with the option to export all these details and exclude self-citations.

3.3 Workplan

The work plan is firstly to set the goals and the features of this software and of course meeting deadlines for each feature. The software is written in java that provides us with more tools in coding.

Firstly, we have to make our app get connected to scholar Google and handle the cookies.

Secondly, we have to manage a way to save the process of each page of scholar google in order to get all the information.

The next step is to find a proper way to submit the query from our software to google and get the result for mining.

After that, our goal is to retrieve the papers and print them to the user with a good taxonomy in a user friendly interface.

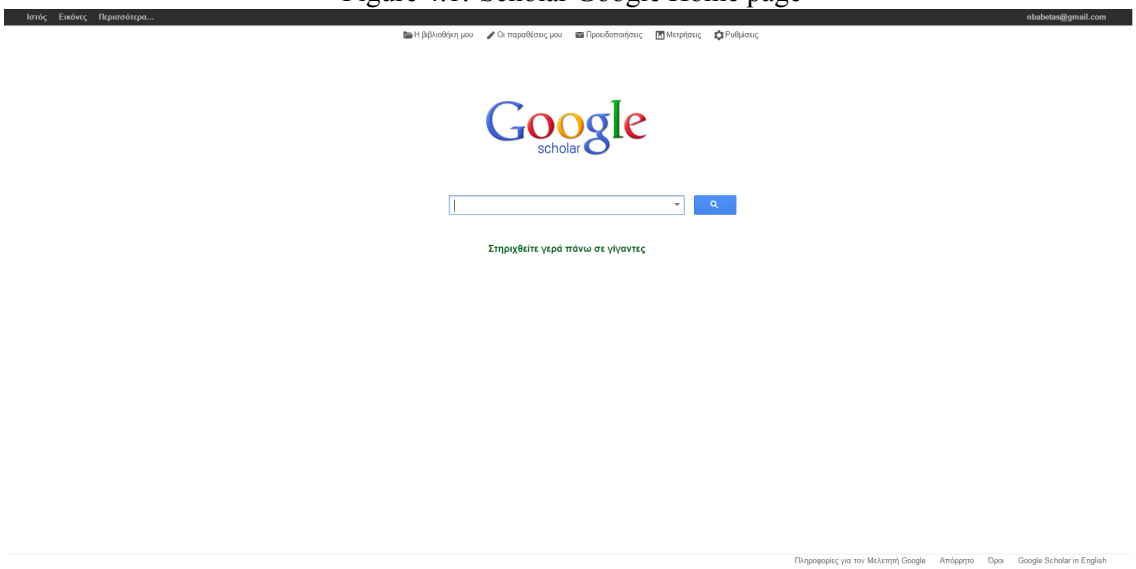
Until now , we only achieve to do things that already other applications do with a great results, so after all we plan to do something new and to retrieve all the citations.

Chapter 4

Data retrieval

Let's see now if someone wants to get a full list of the published papers of someone what he have to do. Firstly he have to connect to Google Scholar with URL : scholar.google.gr. It will give a text field and a button search. In the text field you can put a name of an author after that you can press the button search. Google scholar will return us a full list of his published titles or a user profile and above it the list. If Google has a user profile then we have to get into it because the list that will contains will be more precise and accurate than the simple results as we can see to the below images. **??**

Figure 4.1: Scholar Google Home page



Now that we have a list we can save it by hand title by title and page by page. If we do not have a user profile then we have to collect the papers one by and to go to the next page by pressing next, that contains an new list. On the other hand if there is a profile we have to do the same job as before but it does not have pages instead it has a button that writes show more.

If someone need to retrieve and all the citations as well he have to do exact the same procedure. In user profile every papers has at the right side a number that if someone click on this number google redirect him a new list that contains all the published titles that are referenced in the first one. If we do not have a profile then we can get this list of citations by clicking the link "cited by

Figure 4.2: Scholar Google First result

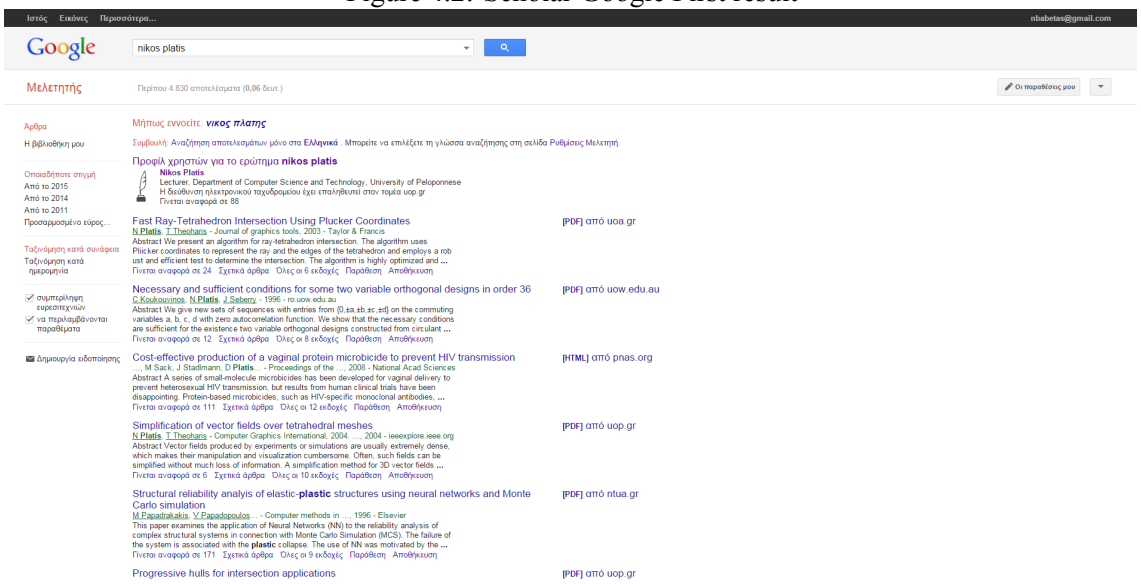
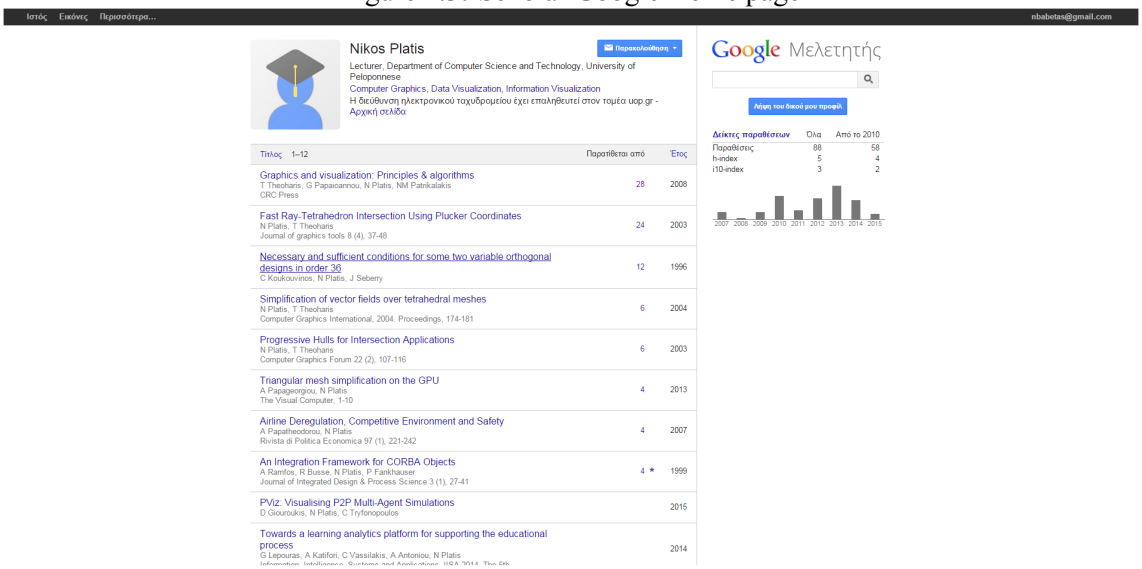


Figure 4.3: Scholar Google Home page



”. When we have the list with all the citations we have to collect the title by title and page by page as before.

Now the user can have all the information that he needed. The time of the retrieval can not be calculated because everything have to be done by hand and if the list is huge then the user will spend weeks in order to retrieve them. He also must separate the the titles that are written from the researcher who we are looking for and he have also to separate the self citation.

Figure 4.4: Scholar Google User profile

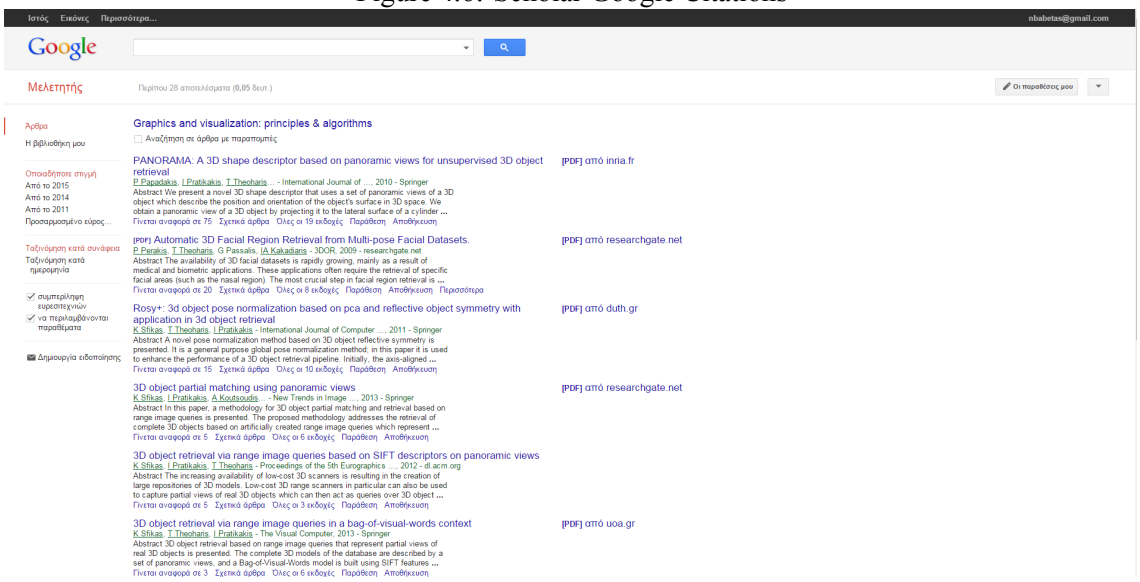
Figure 4.5: Scholar Google User profile

4.1 Nowdays Data retrieval from Google Scholar

We have to notice how a person can make a data retrieval from Google scholar now, without our project. The process is very easy but the collection is the hard part, and of course we have first to check the results if are those that we need.

The user that wants the full title list of a researcher firstly he has to find out how he wrote his name in all the papers. We have to find out how the researcher writes his name because there are a lot of similar names of researchers and this can confuse us and collect titles of an other one.After that we are ready to make our search in google scholar and a list come up to use with all the titles that we were searching for. Now if we want to collect them the only way is to copy and paste the title, publisher and every information that we need. As we can understand this demand a big

Figure 4.6: Scholar Google Citations



amount of to be spend if the researcher has many published papers.

We have to notice now that there are an other option for the user for some researchers. Google Scholar offers profiles for all those that have published papers, which collect by itself all the titles together of this particular researcher. The only requirement is that the researcher would have create a profile. So the user after he find the profile he have just to copy the titles by hand.

But the most difficult of all them is the citation retrieval. As we have collected all the papers the only think that we have to do is to collect all the citations. The searching of the citations is easy as google scholar offers for every paper a link that redirect us to the citation list. This list can be extremely big or small but think if a researcher has 100 papers published and every paper has 100 citations the size of the data is too big to collect by hand. Of course the citations that are offered there are no all of them correct so the user has to check them and delete all those that are wrong. In addition we have to refer that the data collection that will be created it have to be in form that anyone can parse as an xml or bib in order to use the for any purpose.

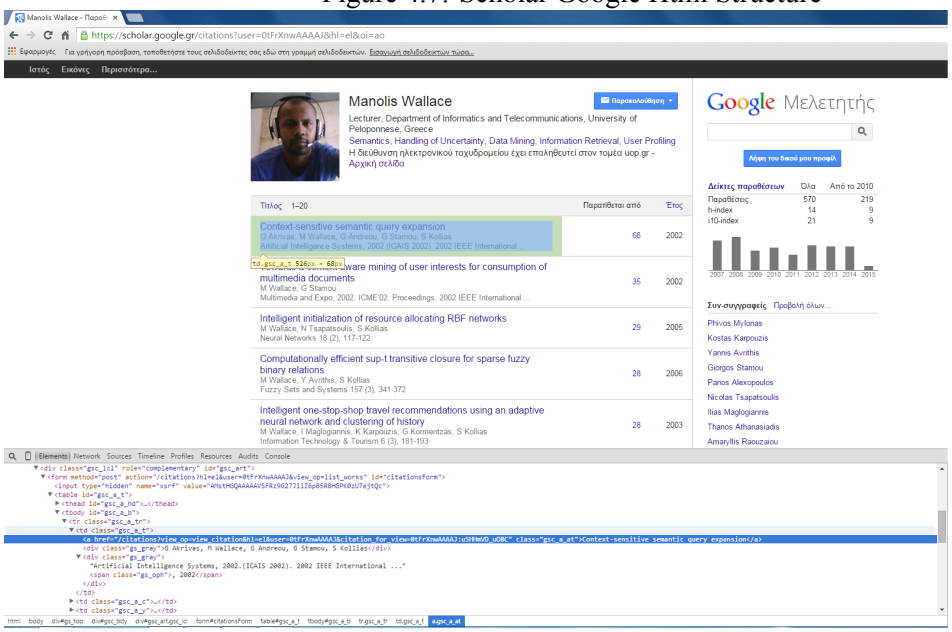
4.2 Publication retrieval structure

Now we have to analyze the Google scholar taxonomy in order to understand what our application has to do. Google scholar offers as we have already mentioned user profiles. User profiles are used, for researchers and provide a full list of their papers, a lot of metrics for the researcher and a link for their citations. But let us see how all this information is projected to the html that our program gets.

Our program can only execute instructions, so the first step is to find the div classes that obtain all the information. In papers it is quite easy as in a div class referred the title, the publisher and the authors of each paper. So we have to split this part of the HTML page and get the information that we need. It is now a simple procedure because there are links and a lot of other functional details that are needed in Google for the display of the information. After hard work we have managed to get only the information that we need. When using a profile there is only one page (one connection)

but if the number of the papers are big we have to press show more papers like a real user.

Figure 4.7: Scholar Google Html Structure



The citation list is quite more difficult because citations are offered on a lot of different pages. In order to parse them our application has to react like a human and press next until it reaches the end of the list. The end of the list is supposedly reached when the next button is disabled. The method in which we retrieve the information of the cited papers is very similar with the one used to retrieve a certain researcher's papers not like to do because we want to automatize this procedure. So to surpass this problem we start saving information after every step and keep a log file as well. When we can reconnect to the google scholar our program will automatically continue to retrieve the information from the point that we have left off. Other ways that we tried to use to overpass the connection problem was to randomize the connections like a real user but we had no luck. But we found a good solution for a quite reconnection and that is to change ip adress.

Lets see how to automate this procedure. First of all, it has to be found the correct url of the researcher. This it is made by taking the full name from the application and combine it with the scholar url. Then we have to find the user profile link if there is one and connect to it in order to retrieve the paper list of the researcher. This procedure is being done from the bellow code segment.

```
CODE SEGMENT THAT FINDS OUT THE LINK OF THE USER PROFILE
    public goToProf(String html) throws MalformedURLException{ // goToProf finds the exact ur
        Pattern pattern;
        pattern = Pattern.compile(”<h4 class=\”gs_rt2\”><a href=\”.*?\”>”);
        Matcher matcher = pattern.matcher(html);
        while (matcher.find()) {
            String[] temp=matcher.group().split(”<a href=\””);
            String[] finalpath=temp[1].split(”>”);
```

```
            url=”http://scholar.google.com”+finalpath[0];


    }
}
    String getURL(){ // getURL returns the url of the of the professor profile in google scho
        return url;
    }
}
```

When our application finds the user profile of the researcher we will retrieve the list of all papers. This can be happened with the connection to the profile, downloading the entire html page and finally by mining this page and keep all the information that are need. The collection from the unprocessed html is made with the help of regular expressions as we can see bellow in the code segment that is called getPapers.

```
CODE SEGMENT THAT REPRESENTS A USER IN ORDER TO COLLECT EVERYTHING
public class getProfile {

    AuthorProfile profile;

    public getProfile(String name) throws IOException, InterruptedException {
        Connection con;
        Save tofile;

        queryFix query = new queryFix(name);// fix researcher name-surname
        String[] search = query.getString();
        URL myURL = new URL(”http://scholar.google.com/scholar?start=0&q=:” + search[0] + ”+a
        con = new Connection(myURL);//connection to scholar google
        goToProf prof = new goToProf(con.getHtml());//Find the url with the list of the paper
        profile = new AuthorProfile();
        int pages = -100;
        do {//collecting all the papers titles
            pages = pages + 100;
            String url = prof.getURL();
            URL u = new URL(url + ”&cstart=” + Integer.toString(pages) + ”&pagesize=100”);//n
            con = new Connection(u);

            if (!con.getHtml().contains(”There are no articles in this profile.”)) {
                getPapers pap = new getPapers(con.getHtml(), search, profile);
            }

        } while (!con.getHtml().contains(”There are no articles in this profile.”));//finish

    }

    AuthorProfile getPro() {
```

```
        return profile;
    }
}
```

Code segment that split cuts and collect the informations from the downloaded Html page

```
public class getPapers {

    AuthorProfile prof;

    public getPapers(String html, String[] name, AuthorProfile prof) {
        this.prof = prof;
        prof.setName(name[0]+name[1]);
        String[] ht = html.split("<tr class=\"gsc_a_tr\">");//split the html page at the poin
        for (int i = 1; i < ht.length; i++) {//loop for all the papers
            Pattern title;
            String[] finalpath = null;
            title = Pattern.compile("<td class=\"gsc_a_t\"><a href=\".*?</a>");//gets title
            Matcher matcher = title.matcher(ht[i]);
            while (matcher.find()) {
                String[] temp = matcher.group().split("class=\"gsc_a_at\">");
                finalpath = temp[1].split("</a>");
            }
            prof.addTitle(finalpath[0]);

            Pattern authors;
            ArrayList<String> a = new ArrayList<>();
            authors = Pattern.compile("</a><div class=\"gs_gray\">.*?</div>");// gets authors
            Matcher matcher2 = authors.matcher(ht[i]);
            finalpath[0]="";
            while (matcher2.find()) {
                if (matcher2.group().compareTo("</a><div class=\"gs_gray\"></div>") != 0) {
                    String[] temp = matcher2.group().split("<div class=\"gs_gray\">");
                    finalpath = temp[1].split("</div>");
                    String[] author = finalpath[0].split(",");
                    for (int j = 0; j < author.length; j++) {
                        a.add(author[j]);

                    }
                }

            }
            prof.addAuthor(a);
            a = new ArrayList<>();

            Pattern journal;
            journal = Pattern.compile("</div><div class=\"gs_gray\">.*?</div>");//get journal
```

```
            Matcher matcher3 = journal.matcher(ht[i]);
            finalpath = null;
            while (matcher3.find()) {
                String[] temp = matcher3.group().split("<div class=\"gs_gray\">");
                finalpath = temp[1].split("<");
                finalpath[0] = finalpath[0].replace(",", ".");
            }
            prof.addJournal(finalpath[0]);

            Pattern cites;
            cites = Pattern.compile("<td class=\"gsc_a_c\"><a href=\".*?</td>");//gets link f
            Matcher matcher4 = cites.matcher(ht[i]);
            finalpath[0]="";
            while (matcher4.find()) {
                if (matcher4.group().compareTo("<td class=\"gsc_a_c\"><a href=\"\" class=\"gs
                    String[] temp = matcher4.group().split("cites=");
                    finalpath = temp[1].split("\"");
                }
            }

            prof.addCite(finalpath[0]);

            Pattern year;
            year = Pattern.compile("<span class=\"gsc_a_h\">.*?</span>");// gets year
            Matcher matcher5 = year.matcher(ht[i]);
            finalpath[0]="";
            while (matcher5.find()) {
                if (matcher5.group().compareTo("<span class=\"gsc_a_h\"></span>") != 0) {
                    String[] temp = matcher5.group().split("<span class=\"gsc_a_h\">");
                    finalpath = temp[1].split("</span>");
                }
            }

            prof.addYear(finalpath[0]);
        }
    }

    AuthorProfile getAuthor() {//returs all the informations for the researcher
        return prof;
    }
}
```

To be more specific if we run all the above, will get an HTML page of google scholar which contains everything we need. Our application succesfully process this HTML and for the nikos platis it can export the useful data like we can see bellow.

TITLE

```
AUTHORS
JOURNAL
YEAR
CITATION LINK
Graphics and visualization: Principles &amp; algorithms
[T Theoharis,  G Papaioannou,  N Platis,  NM Patrikalakis]
CRC Press
2008
16032545112465160752
Fast Ray-Tetrahedron Intersection Using Plucker Coordinates
[N Platis,  T Theoharis]
Journal of graphics tools 8 (4). 37-48
2003
10091455960493275366
Necessary and sufficient conditions for some two variable orthogonal designs in order 36
[C Koukouvinos,  N Platis,  J Seberry]

1996
12695854445630006063
Simplification of vector fields over tetrahedral meshes
[N Platis,  T Theoharis]
Computer Graphics International. 2004. Proceedings. 174-181
2004
12466979524624805722
Progressive Hulls for Intersection Applications
[N Platis,  T Theoharis]
Computer Graphics Forum 22 (2). 107-116
2003
2843585480742285158
Triangular mesh simplification on the GPU
[A Papageorgiou,  N Platis]
The Visual Computer. 1-10
2013
15920707966363123559
Airline Deregulation, Competitive Environment and Safety
[A Papatheodorou,  N Platis]
Rivista di Politica Economica 97 (1). 221-242
2007
1415782540211590679
An Integration Framework for CORBA Objects
[A Ramfos,  R Busse,  N Platis,  P Fankhauser]
Journal of Integrated Design &amp; Process Science 3 (1). 27-41
1999
10661828752338097435,4481689142745147251
Towards a learning analytics platform for supporting the educational process
```

```
[G Lepouras,  A Katifori,  C Vassilakis,  A Antoniou,  N Platis]
Information. Intelligence. Systems and Applications. IISA 2014. The 5th ...
2014
7405226668593159076
PViz: Visualising P2P Multi-Agent Simulations
[D Giouroukis,  N Platis,  C Tryfonopoulos]
Proceedings of the 2015 International Conference on Autonomous Agents and ...
2015

GraphViz and C++-The Boost Graph Library (BGL) is a great way to use Graph Viz to draw directe
[N Platis,  M Tsoukalos]
C/C++ Users Journal. 38-43
2005

Multiresolution Techniques for the Simplification of Triangular and Tetrahedral Meshes
[N Platis]
```

4.3 Citation retrieval

We have reached the point where we have collected all the paper information and we have analyzed the citation list that google scholar offers. So now we have to create some algorithms to procced. After the paper retrieval, our software displays a list of the papers. The user can select some of those papers and press the getCites button to retrieve their citations.

To retrieve the citations our program uses the paper name and checks if we have already retrieved citations for this paper. If we have them it just projects them to the user, if not our program uses the link that we had provided in the previous steps creating the URL for the citation lists and then starts retrieving them. After we get connected to this list our program is executes the same instructions that we asked it to use for the researchers profile, with the difference that it has to parse all the pages until it reaches the end. It Is very important to give some importance to the code that handles the end of list. As we have the link for the citations of each paper our work is to create a function that could connect to it and to parse all the page that there are in order to collect all the above. This is done with the code segments bellow.

```
CODE SEGMENT THAT PARSES ALL THE PAGES AND RETRIEVE ALL CITATION INFORMATIONS

public class getCites {
    HashMap<String, ArrayList<CitesDetails>> hash ;
    ArrayList<CitesDetails> cd = new ArrayList<>();
    boolean flag;
    public getCites(AuthorProfile profile,int value,HashMap<String, ArrayList<CitesDetails>> h
            hash=has;
            flag=false;
```

```
            int pg = -10;
            Connection con;
            hash.put(profile.getPapers().get(value), cd);
            String name=profile.getPapers().get(value);
            Save s =new Save();

            do {
                pg = pg + 10;
                s.SaveSteps(profile.getCites().get(value),pg,name);
                URL u = new URL("http://scholar.google.gr/scholar?start=" + Integer.toString(
                con = new Connection(u);
                Cites cites = new Cites(con.getHtml(), cd);
                Save h = new Save();
                h.Savehash(hash);
            } while (!con.getHtml().contains("<button type=\"button\" aria-label=\"Next\" dis
           if((!con.getHtml().contains("<div id=\"gs_n\" role=\"navigation\">") || con.getHtm
                File file = new File("LastStep.txt");
                file.delete();
            }  else{
               flag=true;
            }
            Save h = new Save();
            h.Savehash(hash);
            CitesList c = new CitesList(cd,name);
            c.setVisible(true);



    }

    boolean getFlag(){
        return flag;

    }
}
```

The pages that citations are projected is quite different that the user's profiles one. So in order to mine the html pages of the citations results we use the code segment bellow.

```
public class Cites {

    public Cites(String html, ArrayList<CitesDetails> cd) throws FileNotFoundException, Unsup
        String[] ht = html.split("<div class=\"gs_r\">");
        for (int k = 1; k < ht.length; k++) {

            CitesDetails det = new CitesDetails();
            Pattern papper;
```

```
papper = Pattern.compile(”<h3 class=\”gs_rt\”>.*?</h3>”);
Matcher papper_matcher = papper.matcher(ht[k]);
while (papper_matcher.find()) {

    String[] finaltitle = null;
    String[] temp = null;

    if (papper_matcher.group().contains(”href”)) {
        temp = papper_matcher.group().split(”href”);
    } else {
        temp = papper_matcher.group().split(”</span>”);
    }
    String[] temp2 = temp[1].split(”>”);
    if (temp2[1].contains(”</a”)) {
        finaltitle = temp2[1].split(”</a”);
    } else {
        finaltitle = temp2[1].split(”</h3”);
    }

    //System.out.printf(”%s”, finalpath[0]);
    finaltitle[0]=finaltitle[0].replace(”&hellip;”, ””);
    det.addPapers(finaltitle[0]);
      //System.out.printf(”%s\n”, finaltitle[0]);
}

Pattern details;
details = Pattern.compile(”<div class=\”gs_a\”>.*?</div>”);// splits for every pa
Matcher details_matcher = details.matcher(ht[k]);

while (details_matcher.find()) {
    String[] result = null;
    String[] temp1=null;
    String[] temp2=null;
    String[] temp = null;
    String[] temp3 = null;
    String[] result2 = null;
    String journal = null;
    ArrayList<String> author;
    if (details_matcher.group().contains(”\”gs_a\”><a href”)) {//take the authors
        temp = details_matcher.group().split(”sra\”>”);
        author = new ArrayList<>();
        for (int i = 1; i < temp.length; i++) {
            result = (temp[i].split(”</a>”));
            result[0]=result[0].replace(”&hellip;”, ””);
            author.add(result[0]);
```

```
                if(temp[i].contains(”, <a”)&& i<temp.length-1){
                   temp1 = (temp[i].split(”, <a”));
                  if(temp1[0].contains(”, ”)){
                    temp2=temp1[0].split(”, ”);
                     author.add(temp2[1]);
                  }

                }
            }
            result2=result[1].split(”-”);
            if(result2.length>2){
                journal = result2[1].replace(”</div>”, ””)+””+result2[2].replace(”</d
                    result2[0]=result2[0].replace(”&hellip;”, ””);
                    author.add(result2[0]);

            }
            else if(result2.length>1){
               journal = result2[1].replace(”</div>”, ””);
                    result2[0]=result2[0].replace(”&hellip;”, ””);
                    author.add(result2[0]);

            }
            else{journal=””;}
            det.addAuthors(author);
            journal = result[1].replace(”</div>”, ””);
            journal = journal.replace(”&hellip;”, ””);
            det.addJournal(journal);

        } else if (details_matcher.group().contains(”href”) && !details_matcher.group
            temp = details_matcher.group().split(”gs_a\”>”);



            result = temp[1].split(”<a href”);
            author = new ArrayList<>();
            result[0]=result[0].replace(”&hellip;”, ””);
            author.add(result[0]);
            for (int i = 1; i < result.length; i++) {
                 temp3=result[i].split(”sra\”>”);
                 result2 = temp3[1].split(”</a>”);
                 result2[0]=result2[0].replace(”&hellip;”, ””);
                 author.add(result2[0]);
            }

            result = result2[1].split(”-”);
```

```
                    if(result.length>2){
                        journal = result[1].replace(”</div>”, ””)+””+result[2].replace(”</div

                    }
                    else if(result.length>1){
                       journal = result[1].replace(”</div>”, ””);

                    }else{journal=””;}
                    journal = journal.replace(”&hellip;”, ””);
                    det.addJournal(journal);
                    det.addAuthors(author);

                } else if (!details_matcher.group().contains(”href”) && !details_matcher.grou
                    temp = details_matcher.group().split(”gs_a\”>”);
                    author = new ArrayList<>();
                    temp3 = temp[1].split(”</div>”);
                    result = temp3[0].split(”-”);
                    result2 = result[0].split(”,”);

                    for (int l = 0; l < result2.length; l++) {
                        //det.addAuthors(result2[l]);
                        result2[l]=result2[l].replace(”&hellip;”, ””);
                        author.add(result2[l]);
                    }
                    if(result.length>2){
                        journal = result[1].replace(”</div>”, ””)+””+result[2].replace(”</div
                    }
                    else if(result.length>1){
                       journal = result[1].replace(”</div>”, ””);
                    }else{
                       journal=””;
                    }
                    journal = journal.replace(”&hellip;”, ””);
                    det.addJournal(journal);
                    det.addAuthors(author);

                }

            }
            cd.add(det);

        }

    }
```

```
}
```

As we refered previously there is a connection limitation from google scholar that can stop us in the middle of our retrieval. In order to avoid any problem we save the information in hash maps and we keep a log of our last step. This can ensure that our data retrieval will not have any loses. When the connection re-established our software continues automatically from last step.

The last step is kept in a file at every single circle of our application. If everything are fine then our application deletes this document.

Chapter 5

Data processing

Data processing is the procedure that should follow retrieval. When we have retrieved the entire html page we have to extract the information which is really needed. In order to have a successful data processing our program must be able to dispose of any trash information and preserve only relevant information. It replaces some html characters that we had retrieved in previous steps. It also uses trim on every string in order to have accurate information. A lot of titles of papers and author names are used multiple times for different procedures so it is very important to have the names as we want them. Another thing that will help us in our data procedure is the perfectly created data structures which our program uses. So after the retrieval and all the trims and replacements of bad characters that are retrieved with the details we use 2 hash maps and 2 custom made data structures that can store all the information of a paper. The 2 hash maps are identical in definition but are used for different reasons. The first one has the name of the researcher as a key and as a result a full list of the papers that he has written. The second one keeps all the citation and has the title of a paper as a key and as a result all his citations. Moreover the two custom made data structures are used in order to store all the details separately in order to use them as we need in the future. So we have categorized in a way all the information that we retrieve and we can use for export.

5.1 Self citations

Self-citation is a reference in a paper that was written by the same author as the one that is being cited. Sometimes we do not want have self-citations because it can create wrong results for researchers. This is why our program offers the option of a document with a full citation list except for self-citations.

When data processing has reached the end, we have to offer the users one more option Our program is going to provide the option of differentiating the list of citations. When have collected everything we will offer an xml document with all the information as an output. Our software will export an xml document which will include everything including self-citations but with distinct separation of them. This will be feasible by interrupting the export of the papers details when they have the same author with the previously searched paper. We keep the names of all the authors of the first paper, and when we process each citation if it has one similar author we tag it as self-citation.

Our idea is to collect everything and after that to create an xml document. In order to do that we send the 2 hash maps in the export class and with the following data segment we create our xml

document.

```
public class Export {

    JFrame frame = new JFrame();

    public Export(String name, HashMap<String, AuthorProfile> hashpro, HashMap<String, ArrayL
        File file = new File(name + ".xml");
        FileWriter fileWriter = new FileWriter(file);
        ArrayList<String> aut;
        try (BufferedWriter print_line = new BufferedWriter(fileWriter)) {
            print_line.write("<?xml version=\"1.0\" encoding=\"UTF-8\"?>" + "\n");
            print_line.write("<research>" + "\n");
            print_line.flush();
            AuthorProfile temp = hashpro.get(name);

            for (int i = 0; i < temp.papers.size(); i++) {
                print_line.write("<full>" + "\n");
                print_line.write("<paper>" + "\n");

                print_line.write("<title>" + temp.getPapers().get(i) + "</title>" + "\n");
                print_line.flush();
                aut = new ArrayList<>();
                for (int k = 0; k < temp.authors.get(i).size(); k++) {
                    String te = temp.authors.get(i).get(k);
                    aut.add(te);
                    print_line.write("<author>" + te + "</author>" + "\n");
                }
                print_line.write("<journal>" + temp.journal.get(i) + "</journal>" + "\n");
                print_line.write("<year>" + temp.year.get(i) + "</year>" + "\n");
                ArrayList<CitesDetails> h = hash.get(temp.getPapers().get(i));
                print_line.write("</paper>" + "\n");

                print_line.write("<citations>" + "\n");
                if (h != null) {
                    int self = 0;
                    for (CitesDetails h1 : h) {
                        if (!h1.pappers.contains("span")) {
                            for (int k = 0; k <= 0; k++) {
                                for (String te : h1.authors) {
                                    if (!aut.isEmpty()) {
                                        if (te.equals(aut.get(k))) {
                                            self++;
                                        }
                                    }
                                }
```

```
                        }

                        if (self > 0) {
                            print_line.write(”<self-citation>” + ”\n”);
                        }
                        print_line.write(”<paper>” + ”\n”);
                        print_line.write(”<title>” + h1.pappers + ”</title>” + ”\n”);
                        print_line.flush();
                        for (String te : h1.authors) {

                            print_line.write(”<author>” + te + ”</author>” + ”\n”);
                        }
                        print_line.write(”<journal>” + h1.journal + ”</journal>” + ”\n”);
                        print_line.write(”</paper>” + ”\n”);
                        if (self > 0) {
                            print_line.write(”</self-citation>” + ”\n”);
                        }
                        self = 0;
                    }
                }
            }

            print_line.write(”</citations>” + ”\n”);
            print_line.write(”</full>” + ”\n”);
        }

        print_line.write(”</research>” + ”\n”);
        JOptionPane.showMessageDialog(frame, ”Export is completed!!!”);
    }

  }
}
```

5.2 Data storage methodology

Our data storage will be done in a xml document which will have the element <full> which will give the title of a paper,the authors, the publisher and all the citations.The information will be given tags each author will be given an <author> tag, the publisher in a <journal> tag and of course the title will be under the <title>. tag The xml document also has a <citation> tag that will include all citations in the paper in question.

Chapter 6

Application

Now the application is ready it can offer a user friendly interface. On the right hand side of the application there is a text field, where the user can type in the researcher's name, and two buttons Search and View Cites.So the user can type the researcher's name and then press search leading to the results being displayed on the left hand side of the application in a list.

When the user has acquired this list he can retrieve the citations of each paper or some of them. One can select the paper or papers that one wants the citations of and after that one can press the view cites button to enable the application to pop lists with the citations. When one has all the information that one needs in the application one can export them in the form of an xml document.

In order to search you have just to put the name of the researcher in the text field and press search. Then the list pannel will be filled with the list of the papers of the researcher. For the citation retrieval the user have to select one or more papers and press view citations.Finally if someone needs to export all these information he has only to select file and then export and an xml document will be created with all the information of the researcher.

There is a possibility to have problem with our connection in google scholar. If something like that happens a message will be inform us that the procedure is paused and it will be continued from the stage that we stopped.Finally the retrieval will be completed when we will get the citation of the last paper.

6.1 Limitations

To sum up we have created a software that behaves like a real user in order to retrieve a precise and accurate list in a very small amount of time. During this procedure we faced some problems in our project that are very important to notice.

The connection to google scholar requires a function that can handle cookies otherwise you can not establish a connection with google. cookie handling was dealt with by getting 2 classes from an other citation software that is offer his code the JabRef.

After we overpassing this problem we tried to get the papers and everything was fine.In citation retrieval we faced our biggest challenge in our project and this was that the after 30 connections can realize that we use a machine and gives use a captcha to solve that we do.

Some ways that we tried to use to overpass the connection problem was to randomize the connections like a real user but we had no luck. But we found a good solution for a quite reconnection and that is to change ip adress.

Figure 6.1: Google scholar retrieval researchers papers

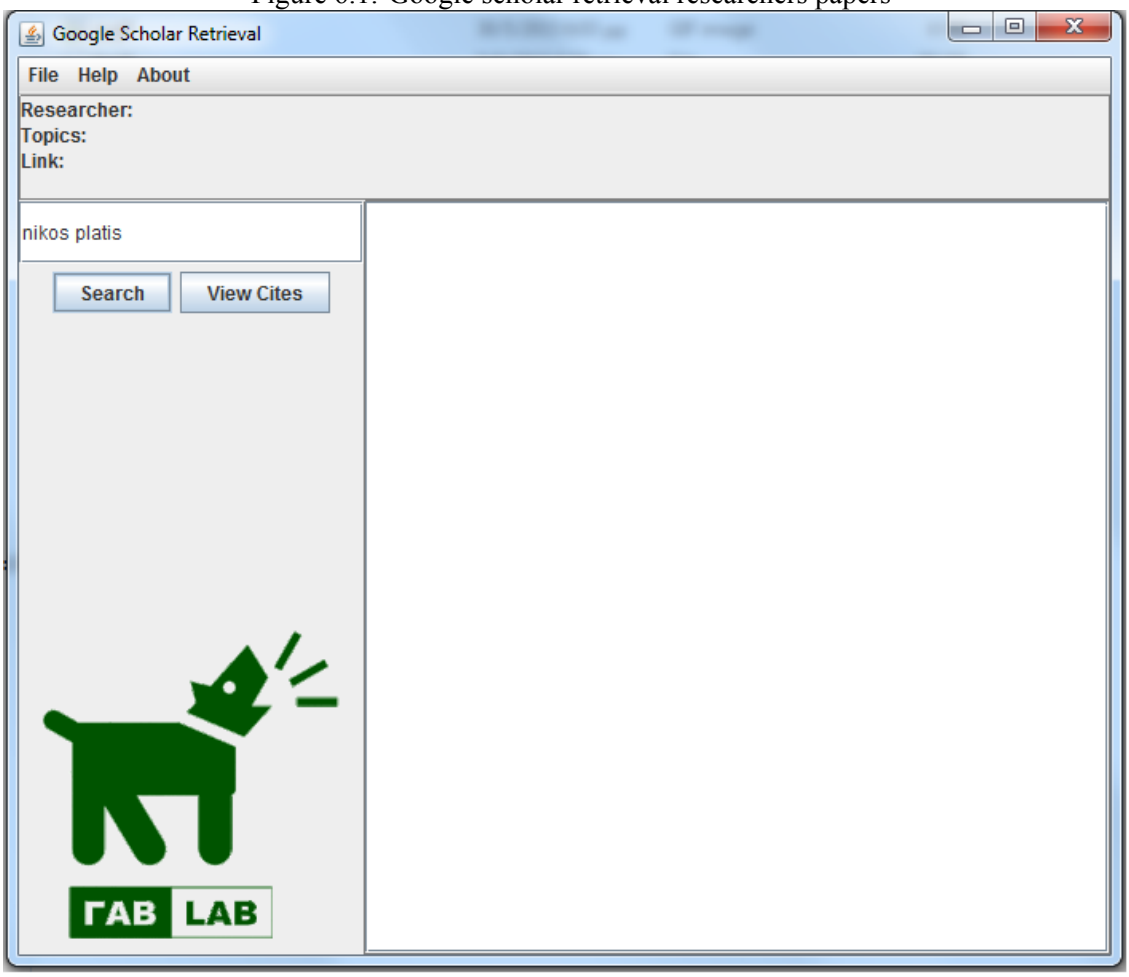


Figure 6.2: Google scholar retrieval citation retrieval

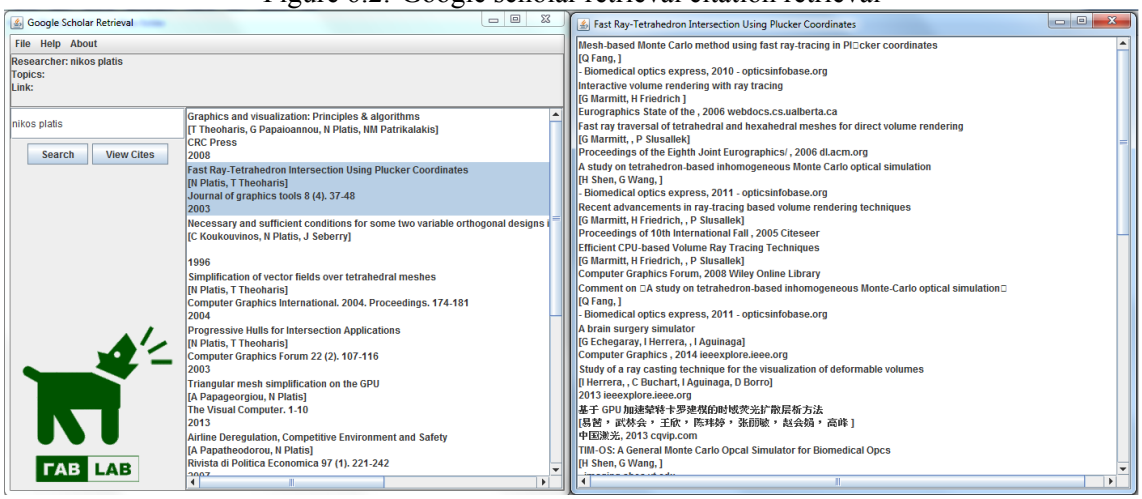


Figure 6.3: Google scholar retrieval Export

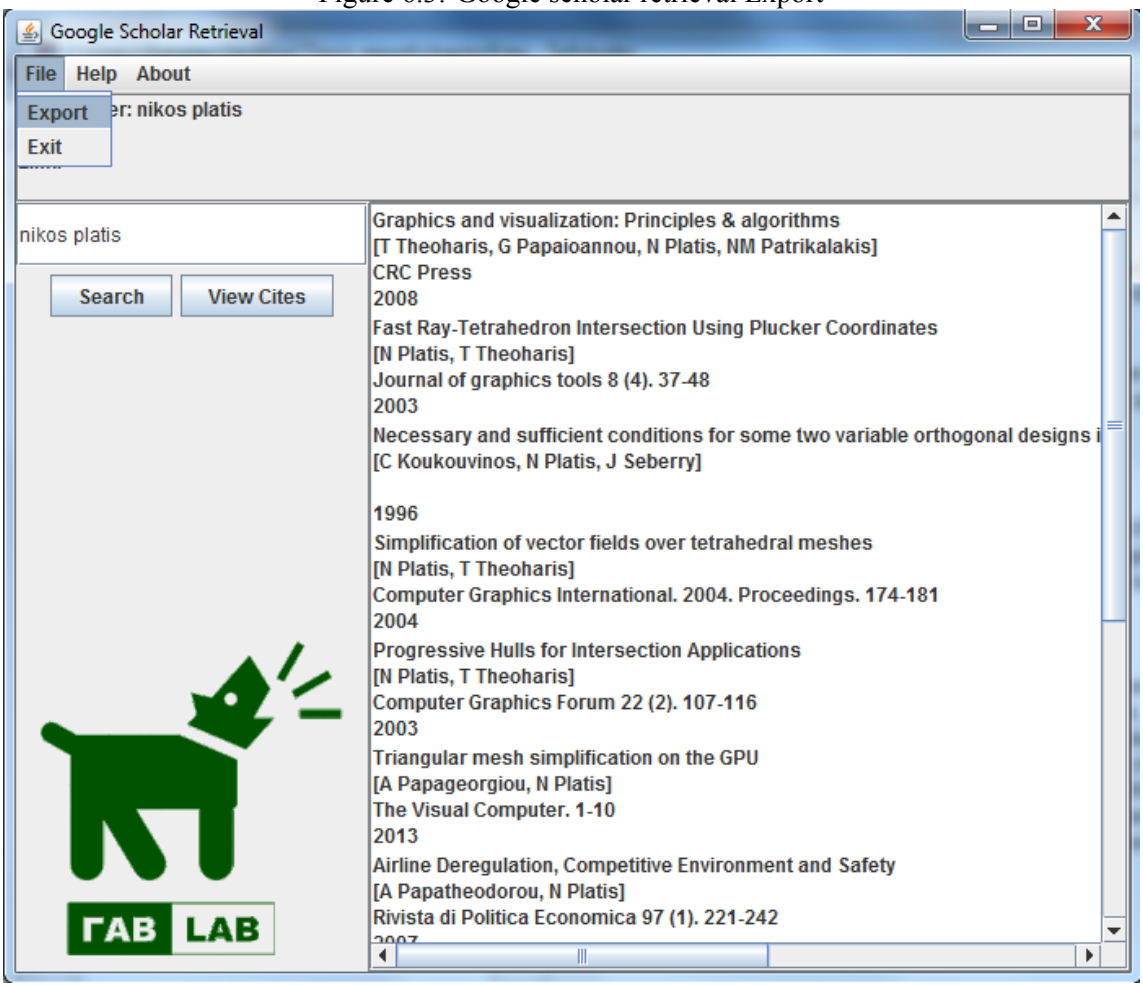
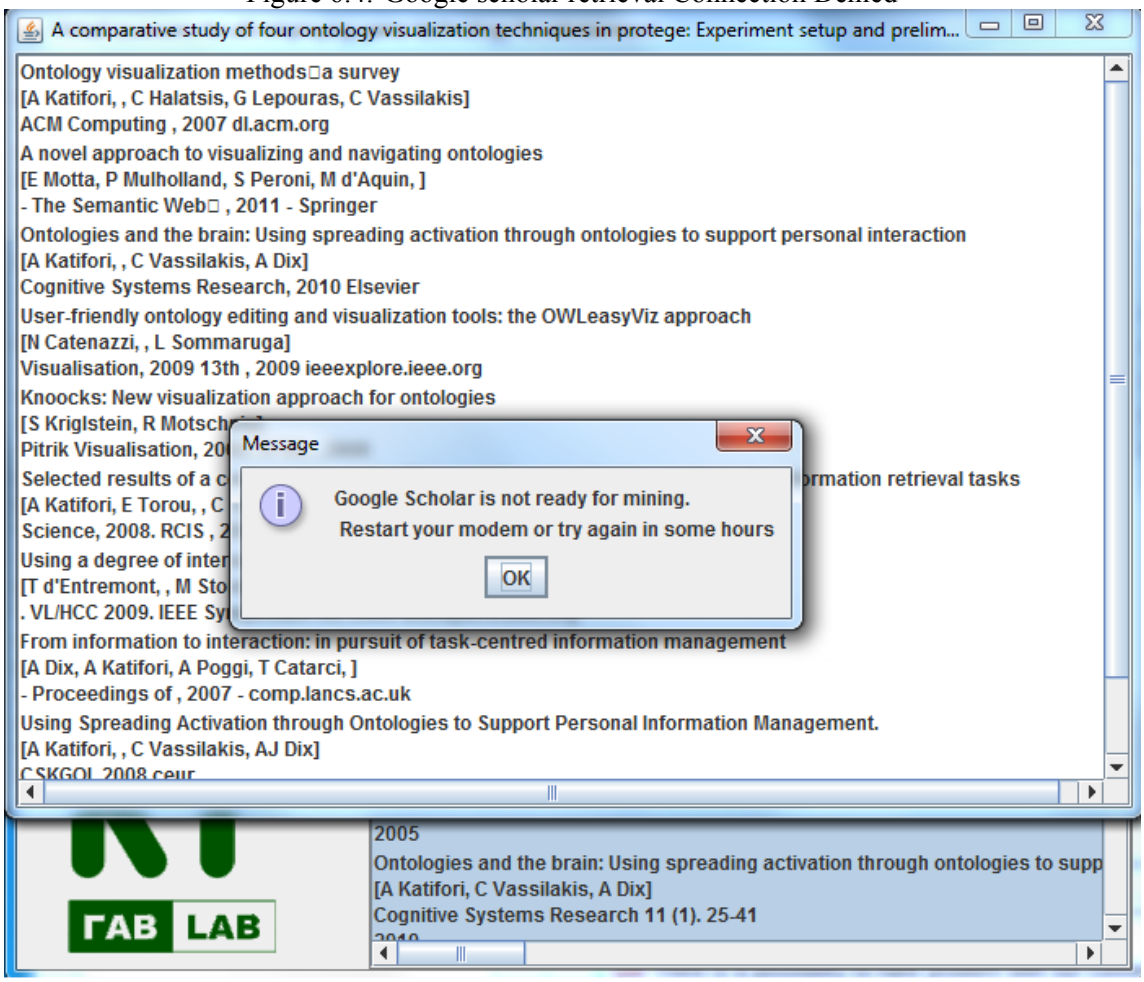


Figure 6.4: Google scholar retrieval Connection Denied



Chapter 7

Trials and results

In our first versions of our application we faced a lot of problems with the data processing and retrieval. We had some problems in the different structures that HTML page of google scholar has for different types of papers. So every time that we have a problem we add an extra option and we customized our code in order to have precise data. Finally we did it and with can collect all the lists that a researcher needs in very small amount of time.

The retrieval of dr Nikos Platis list's needs just 25 sec to collect everything. On the other hand for dr Costas Vassilakis we need about one minute and half without to compute the breaks until google to allow us to get connected to it.

The procedure that our application offers to the public is the fastest because to collect manually all the above list you need to spend some weeks of full work time for the same results as our application.

The final xml of dr Nikos Platis can be found on the appendix.

Chapter 8

Conclusions

In conclusion, our application is ready. We have created an application which can collect a full list of published papers and their citations of a particular researcher. We have handled all problems that we were faced with, the most difficult being the connection denial that Google scholar presents after a number of connections. The data retrieval has been made in the best possible way and we have all the data as we need it in order to use them for extra procedures in the future like export. Our application is connected to Google scholar, where by giving the name of a researcher and it presents a full list of his work, which later the user has the option to export in the form an xml document. It also has a user interface that displays all the previously retrieved information and has buttons for all the procedures. All the procedures that are offered are based on the good data structures that are used in it. Our application can get a lot of extensions. Our application can collect all the published titles for a researcher but if we search for him again after some time it cannot display any newer titles so someone can make this function. It also can export the information only in xml document and not in any other format. Finally the User interface is very basic and it needs a better design in order to become more user-friendly. Out program can also help in other procedures one of such being computing impact factors of a researcher.

Some possibles extensions on our software can be done in the future. A more user friendly method of citation retrieval can be done as the user has to select everything and press view cites several times in order to collect everything. Another extension is a better user interface that will be more user friendly for users. The most important project is to make our application ”invisible” from google scholar in order to evade the denial of connection to it.

```
<?xml version="1.0" encoding="UTF-8"?>

-<research>

-<full>

-<paper>

<title>Graphics and visualization: Principles & algorithms</title>

<author>T Theoharis</author>

<author> G Papaioannou</author>

<author> N Platis</author>

<author> NM Patrikalakis</author>

<journal>CRC Press</journal>

<year>2008</year>

</paper>

-<citations>

-<self-citation>

-<paper>

<title>PANORAMA: A 3D shape descriptor based on panoramic views for unsupervised 3D object re

<author>P Papadakis</author>

<author>I Pratikakis</author>

<author>T Theoharis</author>

<author> </author>
```

<journal> - International Journal of , 2010 - Springer</journal>

</paper>

</self-citation>

-<self-citation>

-<paper>

<title>Automatic 3D Facial Region Retrieval from Multi-pose Facial Datasets.</title>

<author>P Perakis</author>

<author>T Theoharis</author>

<author>G Passalis</author>

<author>IA Kakadiaris</author>

<author> </author>

<journal> - 3DOR, 2009 - researchgate.net</journal>

</paper>

</self-citation>

-<self-citation>

-<paper>

<title>Rosy+: 3d object pose normalization based on pca and reflective object symmetry with ap

<author>K Sfikas</author>

<author>T Theoharis</author>

<author>I Pratikakis</author>

<author> </author>

```
<journal> - International Journal of Computer , 2011 - Springer</journal>

</paper>

</self-citation>


-<paper>

<title>3D object partial matching using panoramic views</title>

<author>K Sfikas</author>

<author>I Pratikakis</author>

<author>A Koutsoudis</author>

<author> </author>

<journal> - New Trends in Image , 2013 - Springer</journal>

</paper>


-<self-citation>


-<paper>

<title>3D object retrieval via range image queries based on SIFT descriptors on panoramic vie

<author>K Sfikas</author>

<author>I Pratikakis</author>

<author>T Theoharis</author>

<author> </author>

<journal> - Proceedings of the 5th Eurographics , 2012 - dl.acm.org</journal>

</paper>

</self-citation>
```

-<self-citation>

-<paper>

<title>3D object retrieval via range image queries in a bag-of-visual-words context</title>

<author>K Sfikas</author>

<author>T Theoharis</author>

<author>I Pratikakis</author>

<author> </author>

<journal> - The Visual Computer, 2013 - Springer</journal>

</paper>

</self-citation>

-<paper>

<title>glGA: an OpenGL Geometric Application framework for a modern, shader-based computer gra

<author>G Papagiannakis</author>

<author>, P Papanikolaou </author>

<journal>, P Papanikolaou - Eurographics , 2014 - george.papagiannakis.org</journal>

</paper>

-<paper>

<title>HS-Patch: A New Hermite Smart Bicubic Patch Modification</title>

<author>V Skala</author>

<author>, M Smolik, L Karlicek, V Ondracka </author>

<journal>, M Smolik, L Karlicek, V Ondracka - submitted for publication, 2013 - naun.org</jou

</paper>

-<self-citation>

-<paper>

<title>A new user-friendly visual environment for breast MRI data analysis</title>

<author>D Antonios, VA Dimitrios, </author>

<author>T Theoharis</author>

<journal> Computer methods and programs , 2013 Elsevier</journal>

</paper>

</self-citation>

-<paper>

<title>Constraint-free topological mapping and path planning by maxima detection of the kerne

<author>P Papadakis</author>

<author>M Gianni</author>

<author>M Pizzoli</author>

<author>F Pirri</author>

<author> </author>

<journal> - International Conference on , 2012 - hal.inria.fr</journal>

</paper>

-<paper>

<title>S-clip E 2: a new concept of clipping algorithms</title>

```
<author>V Skala</author>

<author> </author>

<journal> - SIGGRAPH Asia 2012 Posters, 2012 - dl.acm.org</journal>

</paper>


-<paper>

<title>Evaluation of Adaptive Subdivision Method on Mobile Device</title>

<author>MSM Rahim</author>

<author>, SAM Isa, A Rehman, T Saba </author>

<journal>, SAM Isa, A Rehman, T Saba - 3D Research, 2013 - Springer</journal>

</paper>


-<paper>

<title>Knot visualization experiments for verifiable molecular movies</title>

<author>J Li</author>

<author> TJ Peters</author>

<author> KE Jordan </author>

<journal> arXiv preprint arXiv:1304.5641, 2013 arxiv.org</journal>

</paper>


-<paper>

<title>Bivariate fractal interpolation surfaces: Theory and applications</title>

<author>V Drakopoulos</author>

<author> P Manousopoulos </author>
```

<journal> International Journal of , 2012 World Scientific</journal>

</paper>

-<paper>

<title>Visualisation of complex flows using texture-based techniques</title>

<author>DJ Warne, J Young, </author>

<author>NA Kelson</author>

<journal> ANZIAM Journal, 2013 journal.austms.org.au</journal>

</paper>

-<paper>

<title>A hybrid method for computing the intersection and tangency points of plane curves</ti

<author>D Christou</author>

<author> A Danelakis</author>

<author> M Mitrouli </author>

<journal> Applied Mathematics and , 2012 Elsevier</journal>

</paper>

-<paper>

<title>Interactive diffuse global illumination discretization methods for dynamic environment

<author>AG Gaitatzes</author>

<author> ΑΓ Γκαϊτατζής </author>

<journal> 2012 lekythos.library.ucy.ac.cy</journal>

</paper>

-<paper>

<title>RENDERING TERRAIN ON MOBILE DEVICES USING ADAPTIVE SUBDIVISION METHOD: DESIGN CONSIDERA

<author>SAM Isa, </author>

<author>MSM Rahim</author>

<journal> icts.if.its.ac.id</journal>

</paper>

-<paper>

<title>Navigated Ultrasound Imaging: 3D Reconstruction of Artery Geometry and Flow</title>

<author>Y Karabiyik </author>

<journal> 2013 diva</journal>

</paper>

-<paper>

<title>New Concepts on Two-Dimensional Data Visualization With Applications in Engineering Ana

<author>PA Simionescu</author>

<author>, M Mehrubeoglu </author>

<journal>, M Mehrubeoglu - Journal of , 2012 - .asmedigitalcollection.asme.org</journal>

</paper>

-<paper>

<title>3D model retrieval based on linear prediction coding in cylindrical and spherical proje

<author>V Mehrdad</author>

<author> H Ebrahimnezhad </author>

<journal> Multimedia Tools and Applications, 2014 Springer</journal>

</paper>

-<paper>

<title>Estimation of Skin Optical Parameters for Real-Time Hyperspectral Imaging Applications

<author>A Bj□rgan </author>

<journal> 2013 diva</journal>

</paper>

-<paper>

<title>Partial matching of 3D cultural heritage objects using panoramic views</title>

<author>K Sfikas</author>

<author>I Pratikakis</author>

<author>A Koutsoudis</author>

<author> </author>

<journal> - Multimedia Tools and , 2014 - Springer</journal>

</paper>

-<paper>

<title>3D Navigation with Six Degrees-of-Freedom using a Multi-Touch Display</title>

<author>FR Ortega</author>

<author> </author>

<journal> - 2014 - franciscoraulortega.com</journal>

```
</paper>

-<paper>
<title>Incremental loading of terrain textures</title>
<author>E Fikkan </author>
<journal> 2013 diva</journal>
</paper>

-<paper>
<title>A New Approach to Line□Sphere and Line□Quadrics Intersection Detection and Computation
<author>V Skala</author>
<author> </author>
<journal> - herakles.zcu.cz</journal>
</paper>

-<paper>
<title>Desarrollo de Videojuegos: Programacion Grafica</title>
<author>C Gonzalez</author>
<author> J Albusac</author>
<author> S Perez </author>
<journal> books.google.com</journal>
</paper>

-<paper>
<title>MSc Course Proposal</title>
```

```
<author>G Papagiannakis </author>

<journal> 2011 csd.uoc.gr</journal>

</paper>

</citations>

</full>

-<full>

-<paper>

<title>Fast Ray-Tetrahedron Intersection Using Plucker Coordinates</title>

<author>N Platis</author>

<author> T Theoharis</author>

<journal>Journal of graphics tools 8 (4). 37-48</journal>

<year>2003</year>

</paper>

-<citations>

-<paper>

<title>Mesh-based Monte Carlo method using fast ray-tracing in Pl□cker coordinates</title>

<author>Q Fang</author>

<author> </author>

<journal> - Biomedical optics express, 2010 - opticsinfobase.org</journal>

</paper>
```

-<paper>

<title>Interactive volume rendering with ray tracing</title>

<author>G Marmitt</author>

<author> H Friedrich </author>

<journal> Eurographics State of the , 2006 webdocs.cs.ualberta.ca</journal>

</paper>

-<paper>

<title>Fast ray traversal of tetrahedral and hexahedral meshes for direct volume rendering</t

<author>G Marmitt, </author>

<author>P Slusallek</author>

<journal> Proceedings of the Eighth Joint Eurographics/ , 2006 dl.acm.org</journal>

</paper>

-<paper>

<title>A study on tetrahedron-based inhomogeneous Monte Carlo optical simulation</title>

<author>H Shen</author>

<author>G Wang</author>

<author> </author>

<journal> - Biomedical optics express, 2011 - opticsinfobase.org</journal>

</paper>

-<paper>

<title>Recent advancements in ray-tracing based volume rendering techniques</title>

<author>G Marmitt, H Friedrich, </author>

<author>P Slusallek</author>

<journal> Proceedings of 10th International Fall , 2005 Citeseer</journal>

</paper>

-<paper>

<title>Efficient CPU-based Volume Ray Tracing Techniques</title>

<author>G Marmitt, H Friedrich, </author>

<author>P Slusallek</author>

<journal> Computer Graphics Forum, 2008 Wiley Online Library</journal>

</paper>

-<paper>

<title>Comment on A study on tetrahedron-based inhomogeneous Monte-Carlo optical simulation

<author>Q Fang</author>

<author> </author>

<journal> - Biomedical optics express, 2011 - opticsinfobase.org</journal>

</paper>

-<paper>

<title>A brain surgery simulator</title>

<author>G Echegaray, I Herrera, </author>

<author>I Aguinaga</author>

<journal> Computer Graphics , 2014 ieeexplore.ieee.org</journal>

```
</paper>

-<paper>
<title>Study of a ray casting technique for the visualization of deformable volumes</title>
<author>I Herrera, </author>
<author>C Buchart</author>
<author>I Aguinaga</author>
<author>D Borro</author>
<journal> 2013 ieeexplore.ieee.org</journal>
</paper>

-<paper>
<title>□□ GPU □□□□□□□□□□□□□□□□□□□□□</title>
<author>□□□ □□□□ □□□ □□□□ □□□□ □□□□ □□ </author>
<journal> □□□□, 2013 cqvip.com</journal>
</paper>

-<paper>
<title>TIM-OS: A General Monte Carlo Opcal Simulator for Biomedical Opcs</title>
<author>H Shen</author>
<author>G Wang</author>
<author> </author>
<journal> - imaging.sbes.vt.edu</journal>
</paper>
```

-<paper>

<title>New approach for absolute fluence distribution calculations in Monte Carlo simulations

<author>C B□cklin, D Baumann, </author>

<author>J Fr□hlich</author>

<journal> Journal of Applied Physics, 2014 scitation.aip.org</journal>

</paper>

-<paper>

<title>Fast Ray Traversal of Tetrahedral and Hexahedral Meshes for Direct Volume Rendering</t

<author>T Ertl</author>

<author> K Joy</author>

<author> B Santos </author>

<journal> graphics.cg.unisaarland.de</journal>

</paper>

-<paper>

<title>Volumetric Visualization Techniques of Rigid and Deformable Models for Surgery Simulat

<author>IH Asteasu </author>

<journal> 2013 diglib.eg.org</journal>

</paper>

-<paper>

<title>Interactive volume ray tracing</title>

```
<author>G Marmitt </author>

<journal> 2009 scidok.sulb.uni</journal>

</paper>


-<paper>

<title>Technical Report Fast Ray Traversal of Unstructured Volume Data using Plucker Tests</t

<author>G Marmitt, </author>

<author>P Slusallek</author>

<journal> graphics.cs.unisb.de</journal>

</paper>


-<paper>

<title>Robust Tracking and Advanced Geometry for Monte Carlo Radiation Transport</title>

<author>BM Smith </author>

<journal> 2011 icf4.neep.wisc.edu</journal>

</paper>


-<paper>

<title>A Full 3D GPU-Based Beam Tracing Method for Complex Indoor Environments Propagation Mo

<author>J Tan</author>

<author> Z Su</author>

<author> Y Long </author>

<journal> ieeexplore.ieee.org</journal>

</paper>
```

-<paper>

<title>[C]</title>

<author>T Ernst </author>

<journal> 2009</journal>

</paper>

-<paper>

<title>РАЗРАБОТКА ГЕНЕРАТОРА ДЕКАРТОВЫХ СЕТОК CFD-МОДУЛЯ НА ОСНОВЕ МЕТОДА ВЛОЖЕННЫХ ГРАНИЦ</t

<author>СС Чепилко</author>

<author> ЮВ Юдов </author>

<journal> gidropress.podolsk.ru</journal>

</paper>

-<paper>

<title>Modelling light propagation in Tissue</title>

<author>C B□cklin </author>

<journal> 2014 e</journal>

</paper>

-<paper>

<title>” TOR VERGATA</title>

<author>MG Esposito </author>

<journal> art.torvergata.it</journal>

</paper>

-<paper>

<title>I Simulatori in realt virtuale: un ausilio nella formazione chirurgica</title>

<author>MG Esposito </author>

<journal> 2010 art.torvergata.it</journal>

</paper>

-<paper>

<title>[C]</title>

<author>N Di Lorenzo</author>

<author> AL Gaspari</author>

<author> MG Esposito </author>

<journal> 2010</journal>

</paper>

</citations>

</full>

-<full>

-<paper>

<title>Necessary and sufficient conditions for some two variable orthogonal designs in order

<author>C Koukouvinos</author>

<author> N Platis</author>

<author> J Seberry

```
<journal/>
<year>1996</year>
</paper>

-<citations>

-<paper>
<title>Handbook of combinatorial designs</title>
<author>CJ Colbourn, </author>
<author>JH Dinitz</author>
<journal> 2006 books.google.com</journal>
</paper>

-<self-citation>

-<paper>
<title>Hadamard matrices, orthogonal designs and construction algorithms</title>
<author>S Georgiou</author>
<author>C Koukouvinos</author>
<author>J Seberry</author>
<author> </author>
<journal> - Designs 2002, 2003 - Springer</journal>
</paper>
</self-citation>
```

-<self-citation>

-<paper>

<title>New orthogonal designs and sequences with two and three variables in order 28</title>

<author>C Koukouvinos</author>

<author>J Seberry</author>

<author> </author>

<journal> - Faculty of Informatics-Papers, 2000 - ro.uow.edu.au</journal>

</paper>

</self-citation>

-<self-citation>

-<paper>

<title>Some new three and four variable orthogonal designs in order 36</title>

<author>C Koukouvinos</author>

<author> </author>

<journal> - Journal of statistical planning and inference, 1998 - Elsevier</journal>

</paper>

</self-citation>

-<paper>

<title>New two-variable full orthogonal designs and related experiments with linear regressio

<author>SD Georgiou</author>

<author> </author>

<journal> - Statistics & probability letters, 2007 - Elsevier</journal>

</paper>

-<paper>

<title>When the necessary conditions are not sufficient: sequences with zero autocorrelation

<author>J Horton, </author>

<author>J Seberry</author>

<journal> 1999 ro.uow.edu.au</journal>

</paper>

-<self-citation>

-<paper>

<title>Necessary and sufficient conditions for three and four variable orthogonal designs in

<author>S Georgiou</author>

<author>C Koukouvinos</author>

<author>, M Mitrouli </author>

<journal>, M Mitrouli - Journal of statistical , 2002 - Elsevier</journal>

</paper>

</self-citation>

-<paper>

<title>When the necessary conditions are not suficient: sequences with zero autocorrelation f

<author>J rey Horton, </author>

```
<author>J Seberry</author>

<journal> Citeseer</journal>

</paper>


-<paper>

<title>[C]</title>

<author>J Horton, </author>

<author>J Seberry</author>

<journal> Bulletin of the Institute of Combinatorics and Its , 1999 The Institute</journal>

</paper>


-<paper>

<title>CHARLES J. COLBOURN</title>

<author>JH DINITZ</author>

<author> </author>

<journal> - 1996 - uow.edu.au</journal>

</paper>


-<paper>

<title>Boolean functions, Hadamard matrices, orthogonal designs applicable to security and co

<author>T Xia </author>

<journal> 2001 ro.uow.edu.au</journal>

</paper>


-<paper>
```

<title>31 Orthogonal Designs</title>

<author>J SEBERRY</author>

<author>, R CRAIGEN </author>

<journal>, R CRAIGEN - CRC Handbook of Combinatorial , 2010 - books.google.com</journal>

</paper>

</citations>

</full>

-<full>

-<paper>

<title>Simplification of vector fields over tetrahedral meshes</title>

<author>N Platis</author>

<author> T Theoharis</author>

<journal>Computer Graphics International. 2004. Proceedings. 174-181</journal>

<year>2004</year>

</paper>

-<citations>

-<paper>

<title>A Delaunay simplification algorithm for vector fields</title>

<author>TK Dey</author>

<author>JA Levine</author>

```
<author>, R Wenger </author>

<journal>, R Wenger - Computer Graphics and , 2007 - ieeexplore.ieee.org</journal>

</paper>


-<paper>

<title>□□□□□□□□□□□ LOD □□□□□□□</title>

<author>□□□ </author>

<journal> Computer Engineering and Applications, 2010 cea.ceaj.org</journal>

</paper>


-<paper>

<title>Hi□rarchisation et visualisation multir□solution de r□sultats issus de codes de simula

<author>F Vivodtzev </author>

<journal> 2005 tel.archives</journal>

</paper>


-<self-citation>


-<paper>

<title>Multiresolution Techniques for the Simplification of Triangular and Tetrahedral Meshes

<author>N Platis</author>

<author> </author>

<journal> - cgi.di.uoa.gr</journal>

</paper>

</self-citation>
```

-<paper>

<title>Delaunay Methods for Approximating Geometric Domains</title>

<author>JA Levine</author>

<author> </author>

<journal> - 2009 - rave.ohiolink.edu</journal>

</paper>

-<paper>

<title>□□□□□□□□□□□□□□□□□□□□□□□□</title>

<author>□□ </author>

<journal> □□□□, 2012 cqvip.com</journal>

</paper>

</citations>

</full>

-<full>

-<paper>

<title>Progressive Hulls for Intersection Applications</title>

<author>N Platis

<author> T Theoharis</author>

<journal>Computer Graphics Forum 22 (2). 107-116</journal>

<year>2003</year>

```
</paper>

-<citations>

-<self-citation>

-<paper>
<title>Multiresolution Techniques for the Simplification of Triangular and Tetrahedral Meshes
<author>N Platis</author>
<author> </author>
<journal> - cgi.di.uoa.gr</journal>
</paper>
</self-citation>

-<paper>
<title>Progressive Hulls: Application on Biomedical Data</title>
<author>D Cholt </author>
<journal> cescg.org</journal>
</paper>

-<paper>
<title>” TOR VERGATA</title>
<author>MG Esposito </author>
<journal> art.torvergata.it</journal>
</paper>
```

-<paper>

<title>I Simulatori in realt virtuale: un ausilio nella formazione chirurgica</title>

<author>MG Esposito </author>

<journal> 2010 art.torvergata.it</journal>

</paper>

-<paper>

<title>[C]</title>

<author>N Di Lorenzo</author>

<author> AL Gaspari</author>

<author> MG Esposito </author>

<journal> 2010</journal>

</paper>

-<paper>

<title> cage </title>

<author> </author>

<journal> Journal of Software, 2013 jos.org.cn</journal>

</paper>

</citations>

</full>

-<full>

-<paper>

<title>Triangular mesh simplification on the GPU</title>

<author>A Papageorgiou</author>

<author> N Platis</author>

<journal>The Visual Computer. 1-10</journal>

<year>2013</year>

</paper>

-<citations>

-<paper>

<title>Journal Finder Assignment 91.546Computer Graphics I</title>

<author>K Vyas </author>

<journal> ACM Transactions on Graphics (TOG), 2015 cs.uml.edu</journal>

</paper>

-<paper>

<title>A novel simulation framework based on information asymmetry to evaluate evacuation pla

<author>X Che, Y Niu, B Shui, J Fu, G Fei, </author>

<author>P Goswami</author>

<journal> The Visual , 2015 Springer</journal>

</paper>

-<paper>

<title>Optimizaci n y paralelizaci n de un algoritmo de generaci n de Skeletons a partir de m

<author>IY Rojas Hern□ndez </author>

<journal> 2014 tesis.uchile.cl</journal>

</paper>

-<paper>

<title>Обзор методов упрощения полигональных моделей на графическом процессоре</title>

<author>ВИ Гонахчян </author>

<journal> ispras.ru</journal>

</paper>

</citations>

</full>

-<full>

-<paper>

<title>Airline Deregulation, Competitive Environment and Safety</title>

<author>A Papatheodorou</author>

<author> N Platis</author>

<journal>Rivista di Politica Economica 97 (1). 221-242</journal>

<year>2007</year>

</paper>

-<citations>

-<paper>

<title>The Supply of Tourism Services: Hospitality, Transport, Attractions</title>

<author>G Candela, </author>

<author>P Figini</author>

<journal> The Economics of Tourism Destinations, 2012 Springer</journal>

</paper>

-<paper>

<title>Economic Rationale for Licensing Tour Guides in Poland</title>

<author>A Pawlicz</author>

<author> </author>

<journal> - Tourism in Southern and Eastern Europe, 2013 - papers.ssrn.com</journal>

</paper>

-<paper>

<title>Production in Tourism</title>

<author>G Candela, </author>

<author>P Figini</author>

<journal> The Economics of Tourism Destinations, 2012 Springer</journal>

</paper>

-<paper>

<title>Ekologia jako podstawa innowacyjnego produktu hotelarskiego</title>

<author>A Pawlicz</author>

<author>M Sidorkiewicz</author>

```
<author> </author>

<journal> - 2013 - depot.ceon.pl</journal>

</paper>

</citations>

</full>


-<full>


-<paper>

<title>An Integration Framework for CORBA Objects</title>

<author>A Ramfos</author>

<author> R Busse</author>

<author> N Platis</author>

<author> P Fankhauser</author>

<journal>Journal of Integrated Design & Process Science 3 (1). 27-41</journal>

<year>1999</year>

</paper>


-<citations>


-<paper>

<title>Bio-AXS: Uma Arquitetura para Integra□□o de Fontes de Dados e Aplica□□es de Biologia M

<author>LFB Seibel </author>

<journal> 2000 uniriotec.br</journal>
```

```
</paper>

-<paper>
<title>Using an object reference approach to distributed updates</title>
<author>D Kambur, </author>
<author>M Roantree</author>
<journal> Database and Expert Systems , 2007 Springer</journal>
</paper>

-<paper>
<title>Objektorientierte Sichten zur Integration statisch getypter Objektbanken</title>
<author>R Busse </author>
<journal> 2004 tuprints.ulb.tu</journal>
</paper>

-<paper>
<title>Komposition von XML Dataflows zur Integration heterogener WWW-Informationsdienste</tit
<author>J Henkel, I Macherius, </author>
<author>P Fankhauser</author>
<journal> 2000 www</journal>
</paper>
</citations>
</full>

-<full>
```

-<paper>

<title>PViz: Visualising P2P Multi-Agent Simulations</title>

<author>D Giouroukis</author>

<author> N Platis</author>

<author> C Tryfonopoulos</author>

<journal/>

<year>2015</year>

</paper>

<citations> </citations>

</full>

-<full>

-<paper>

<title>Towards a learning analytics platform for supporting the educational process</title>

<author>G Lepouras</author>

<author> A Katifori</author>

<author> C Vassilakis</author>

<author> A Antoniou</author>

<author> N Platis</author>

<journal>Information. Intelligence. Systems and Applications. IISA 2014. The 5th ...</journal

<year>2014</year>

</paper>

```
<citations> </citations>

</full>

-<full>

-<paper>

<title>GraphViz and C++-The Boost Graph Library (BGL) is a great way to use Graph Viz to draw

<author>N Platis</author>

<author> M Tsoukalos</author>

<journal>C/C++ Users Journal. 38-43</journal>

<year>2005</year>

</paper>

<citations> </citations>

</full>

-<full>

-<paper>

<title>Multiresolution Techniques for the Simplification of Triangular and Tetrahedral Meshes

<author>N Platis</author>

<journal/>

<year/>

</paper>

<citations> </citations>
```

```
</full>

</research>
```

Bibliography

[] *Histogram*. URL: www.cs.armstrong.edu/liang/intro6e/evennumberedexercise/Histogram.java.

[Blo14] Google Scholar Blog. *Google Scholar Citations Open To All*. retrieved 2. Aug. 2014.

[cit] citeseer. *CiteSeer*. URL: http://citeseerx.ist.psu.edu/index;jsessionid=D9E8866A50810F1B528D2F84007C2C09.

[Dan05] Sips M. Daniel A. Keim. "Visual Data-Mining Techniques". In: *The Visualization Handbook* 43 (2005), pp. 831–843.

[goo] google. *Scholar Google*. URL: https://scholar.google.gr/.

[har] harzing. *Publish or Perish*. URL: http://www.harzing.com/pop.htm.

[jab] jabref. *Jabref*. URL: http://jabref.sourceforge.net/.

[Kei02] Daniel A. Keim. "Information Visualization and Visual Data Mining". In: *IEEE transactions on visualization and computer graphics* 7.1 (Mar. 2002), pp. 100–107.

[Mic] Microsoft. *Academic search*. URL: http://academic.research.microsoft.com/.

[Ora] Oracle. *Reading XML Data into a DOM*. URL: https://docs.oracle.com/javase/tutorial/jaxp/dom/readingXML.html.

[ove] Stack overflow. *Triangles and lines*. URL: http://stackoverflow.com/questions/3536428/draw-a-line-at-a-specific-angle-in-java.

[reu] thomson reuters. *Web Of Science*. URL: http://www.webofscience.com.

[sco] scopus. *scopus*. URL: http://www.elsevier.com/solutions/scopus.

[W3s] W3schools. *XML Tree*. URL: http://www.w3schools.com/xml/xml_tree.asp.

[Wal14] Manolis Wallace. "Extracting and visualizing research impact semantics". In: *SMAP 9th International Workshop on Semantic and Social Media Adaptation and Personalization* (Nov. 2014), pp. 104–109.

[Wika] Wikipedia. *Applied sciences*. URL: http://en.wikipedia.org/wiki/Outline_of_applied_science#Fields_of_applied_science.

[Wikb] Wikipedia. *List of academic databases and search engines*. URL: http://en.wikipedia.org/wiki/List_of_academic_databases_and_search_engines.

[Wikc] Wikipedia. *List of academic databases and search engines*. URL: http://en.wikipedia.org/wiki/List_of_academic_databases_and_search_engines.

[Wikd] Wikipedia. *Natural sciences*. URL: http://en.wikipedia.org/wiki/Natural_science.

[Wike] Wikipedia. *Science classifications*. URL: http://el.wikipedia.org/wiki/%C3%8E%C2%95%C3%8F%C2%80%C3%8E%C2%B9%C3%8F%C2%83%C3%8F%C2%84%C3%8E%C2%AE%C3%8E%C2%BC%C3%8E%C2%B7.

[Wikf] Wikipedia. *Sciences/Categories*. URL: http://en.wikipedia.org/wiki/Portal:Science/Categories.

[Wikg] Wikipedia. *Social sciences*. URL: http://en.wikipedia.org/wiki/List_of_social_sciences.

[Wik04] Wikipedia. *List of scientific journals*. retrieved 3. Aug. 2004. URL: http://en.wikipedia.org/wiki/List%20of%20scientific%20journals.

[Wik12] Wikipedia. *ACM Computing Classification System*. 2012. URL: http://www.acm.org/about/class/class/2012.