

## Welcome to Oyster Search

Citation chaining is a strategy employed by academics during a literature search. Chaining can operate backwards or forwards in time. Backward chaining involves browsing through the reference list of a known article of interest to access key previous works. Forward chaining involves using a citation index service, such as Web of Science or Scopus, to locate more recently published articles that have cited known articles.

Backward and forward chaining can be used in combination, resulting in a powerful search method called citation cycling. Typically, one would locate an important article, note its references, and then retrieve all the articles that cite these (or a selection of these) references. One would then retrieve all articles that are referenced by these citing articles. Following this iterative process, it is possible to quickly build a comprehensive picture of the literature that relates to your topic of interest.

Oyster search employs a technique known as Citation Chain Aggregation (CCA), which was conceived to facilitate the process of citation cycling. Current Web interfaces only allow you to chain backwards or forwards from a single article of interest. CCA removes this limitation, by allowing you to iteratively build a list of relevant articles, known as a Pearl list, and immediately see all other articles that either cite or are cited by ALL of the articles in your Pearl list. Most importantly, the more Pearl articles an articles relates to the higher it will be ranked in the cited/citing list.

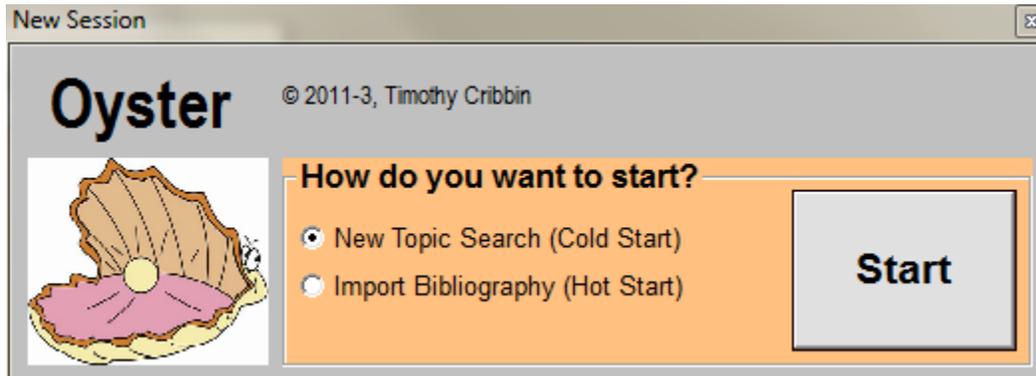
CCA is an exciting concept, but we are keen to see how academic researchers actually make use of it as part of their overall research activity. To this end, we have implemented a prototype tool. We invite you to use this tool, free of charge, as part of your normal research activity. Please feel free to contact us to report bugs, make suggestions or just tell us what you think. You can use the email address shown on the [home page](#) or, if you wish, make a comment on Twitter using the [#oystersearch](#) tag.

We are not sure exactly which search tasks or goals CCA will best serve. However, we have identified three general types of search scenario that CCA might promote. We refer to these throughout this document as prototype task scenarios. These scenarios are described in greater detail later. Please read their descriptions carefully before you start using the tool. If you find a use for CCA that does not fit into these scenario definitions, please let us know.

## Using Oyster

When you first load up Oyster you are presented with a start-up screen as shown below. There are two ways to start a new session. Selecting a New Topic Search, then clicking on the Start button, will take you straight to the main tool interface, where you will begin with an empty Pearl list. Follow this route if you want to begin with a keyword database search. Alternatively, selecting Import Bibliography, followed by Start, will open a file

dialogue box where you can choose an existing list of references to initially populate your Pearl list with.



### ***New Topic Search (Cold Start)***

If you selected a new topic search, then you will need to use the search function to start importing items into the Pearl list. The procedure for Import Bibliography is described later in this document. The figure below shows the interface with the Search tab selected. You can see there are four search fields: Text, Title, Author and Year. Use the Text field when you want to search for terms anywhere within a document.

When you have specified the search fields hit Enter or click the Search button and the tool will retrieve the top 100 search results from the citation database. This version of CCA uses the Microsoft Academic Search<sup>1</sup> (MAS) database. The results are displayed in the adjacent table. Initially they are ordered by relevance, but you can reorder the records according to any of the named columns (e.g. citation count, author name, year etc.) by clicking on the column header. Clicking a second time on the same column header will reverse the order of records.

To help you decide if a search record is relevant, further information can be gained in two ways. First, if you click on a record a tool-tip will appear showing you the first sentence or two of the abstract. Alternatively, if you right-click and select Navigate to Source, then the tool will attempt to open the record on the publisher's web-site. If the location is not known, then the title will be submitted to Google Scholar and the result shown.

---

<sup>1</sup> Microsoft Academic Search: <http://academic.research.microsoft.com/>

The screenshot shows the Oyster search interface. On the left, there is a search form with the text 'cloud computing' entered. The main area displays a search results table with columns: Rel, Cited, Author, Year, Title, Source, PK, DOI, and Abstract. A context menu is open over the fourth row, showing 'Add to Pearl' and 'Navigate to Source' options.

Rel	Cited	Author	Year	Title	Source	PK	DOI	Abstract
001	0463	Armbrust, ...	2009	Above the Clouds: A Berkeley View of Clo...		52...		
002	0258	Nurmi, Da...	2009	The Eucalyptus Open-Source Cloud-Comp...	Cluster Computing and the Grid	51...	10...	Cloud comp...
003	0255	Buyya, Ra...	2009	Cloud Computing and Emerging IT Platform...	Future Generation Computer Systems	50...	10...	
004	0166	Buyya, Ra...	2008	Market-Oriented Cloud Computing: Meas...	High Performance Computing and Commun...	43...	10...	This keywor...
005	0179	Foster, Ian	2008	Cloud Computing and Grid Co...	Environments Workshop	50...	10...	
006	0147	Hayes, Brian	2008	Cloud computing	The ACM	43...	10...	As software
007	0104	Armbrust, ...	2010	A view of cloud computing	The ACM	13...	10...	Clearing the
008	0075	Vouk, Mia...	2008	Cloud Computing - Issues, Research...	International Conference on Information Tec...	39...	10...	IdquoCloudct
009	0073	Youseff, La...	2008	Toward a Unified Ontology of Cloud Comp...	Grid Computing Environments Workshop	55...	10...	
010	0012		1979	A terrain and cloud computer image generat...	ACM Siggraph Computer Graphics	88...	10...	A method is
011	0080	Foster, Ian	2009	Cloud Computing and Grid Computing 360...	Computing Research Repository	47...		
012	0055	Rochweg...	2009	The Reservoir model and architecture for o...	Ibm Journal of Research and Development	51...	10...	The emergin
013	0055	Wang, Lizhe	2008	Scientific Cloud Computing: Early Definitio...	High Performance Computing and Commun...	47...	10...	The Cloud o
014	0057	Buyya, Ra...	2009	Modeling and Simulation of Scalable Cloud...	Computing Research Repository	60...	10...	
015	0053	Frey, Rich...	1999	A comparison of cloud top heights compute...	Journal of Geophysical Research	62...	10...	Data from tw
016	0049	Santos, Nu...	n/a	Towards Trusted Cloud Computing		51...		Cloud compo
017	0044	Grossman, ...	2009	The Case for Cloud Computing	IT Professional	60...	10...	To understa

## Adding an article to the Pearl list

When you find a record that you want to add to the Pearl list, right-click the item and select Add to Pearl from the menu. If you look above the Search box, you will see three near identical list views. The selected record will be added to the Pearl (centre) list. After a few seconds, new records will appear in both the left and the right hand lists. The left-hand list shows the articles that are cited (referenced) by the Pearl article, whilst the right-hand list shows the documents that cite the Pearl article. We will see later on that as further documents are added to the Pearl, these lists will increase in size.

However, for now let's focus on the view after just the first article has been added. In the next screenshot shown on the next page, the article entitled "Market-Oriented Cloud Computing" by Buyya has been added to the Pearl. If you look below the main list views, you will see that the Article Summary tab is now selected and shows detailed metadata, including the full abstract, associated with the article. This display is updated whenever you click on a new record in any of the three main list views.

The Pearl, Cited (left-hand) and Citing (right-hand) lists all display the same metadata about their records. Working from left to right, the first column is perhaps the most important one. It displays a value called the overlap score. This describes the number of Pearl documents that each record relates to and by default all lists are sorted, in descending order, by this column. As there is currently only one Pearl document, the overlap score is 1 for all records on display, except the Pearl itself. This means that all the Citing documents cite the one (and only) Pearl document, whilst all Cited documents are referenced by the one (and only) Pearl document. The Pearl itself has a score of zero because it does not cite itself. Whilst this value is not very informative at the moment, we shall see later how it becomes much more useful as the Pearl list grows.

The second column, Cited, indicates the citation count for each document. This is followed by article title and source (Journal/Conference). PK and GI are the unique key numbers within the MAS and local databases respectively. As with the search list, these lists can be reordered by any column by clicking on the header.

Oyster Search v1.00 beta © Timothy Cribbin, 2013

Session Article View

Cl.	Cited	Author	Year	Title	Source
001	0242	BUYYA, RAJ...	2005	The Grid Economy	Proceeding
001	0118	Kashay, Kate...	2005	Virtual workspaces: Achieving gu...	Scientific P
001	0095	Inwin, David	2006	Sharing Networked Resources w...	USENIX Te
001	0055	AuYoung, Alvin	2004	Resource Allocation in Federated ...	
001	0050	Lai, Kevin	2004	Tycoon: an Implementation of a D...	Computing
001	0049	CHETTY, MA...	2002	Weaving computational grids: ho...	Computing
001	0041	Chu, Xingchen	2007	Aneka: Next-Generation Enterpris...	e-Science :
001	0021	Venugopal, Sr...	2008	A Negotiation Mechanism for Adv...	Internationa

8 articles cited by pearl articles

Cl.	Cited	Author	Year	Title	Source
000	0199	Buyya, Rajku...	2008	Market-Oriented Cloud Computing	High Perform

1 items in the pearl list

Cl.	Cited	Author	Year	Title	Source
001	0261	Buyya, Rajku...	2009	Cloud Computing and Emerging I...	Future
001	0027	Jha, Shantenu	2009	Using clouds to provide grids wit...	Concour
001	0021	Liu, Liang	2009	GreenCloud: a new architecture f...	
001	0016	Briscoe, Gerard	2009	Digital Ecosystems in The Cloud...	Compu
001	0014	Yigibasi, Nezih	2009	C-Meter: A Framework for Perfor...	
001	0014	Matsunaga, A...	2008	CloudBLAST: Combining MapRe...	
001	0013	Anstett, Tobias	2009	Towards BPEL in the Cloud: Exp...	IEEE C
001	0011	Chieu, Trieu	2009	Dynamic Scaling of Web Applicati...	IEEE In
001	0011	Calheiros, Ro...	2009	CloudSim: A Novel Framework f...	Compu
001	0011	Truong, Hong...	2009	On analyzing and specifying con...	Asia-Pt
001	0010	Mennis, Alex...	2009	Community Cloud Computing	
001	0009	Beloglazov, A...	2010	Energy Efficient Resource Manag...	Cluster
001	0009	Ostermann, Si...	2009	Extended Grids with Cloud Reso...	
001	0009	Vecchiola, Ch...	2009	Multi-Objective Problem Solving ...	Compu
001	0008	Hassan, Moh...	2009	A framework of sensor-cloud inte...	Internat
001	0008	Ranjan, Rajiv	2008	Decentralized Overlay for Federat...	Compu
001	0007	Jha, Shantenu	n/a	Using Clouds to Provide Grids Hi...	

198 articles citing pearl articles

Article Summary Search

AUTHORS: Buyya, Rajkumar; Yeo, Chee, Venugopal, Srikumar,  
YEAR: 2008  
TITLE: Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities  
SOURCE: High Performance Computing and Communications  
ARTICLE TYPE: Conference  
ABSTRACT: This keynote paper: presents a 21 st century vision of computing; identifies various computing paradigms promising to deliver the vision of computing utilities; defines Cloud computing and provides the architecture for creating market-oriented Clouds by leveraging technologies such as VMs; provides thoughts on market-based resource management strategies that encompass both customer-driven service management and computational risk management to sustain SLA-  
CITED: 199 times  
ID: 4373636  
DOI: 10.1109/HPCC.2008.172  
URL: http://www.buyya.com/gridbus/papers/hpcc2008\_keynote\_cloudcomputing.pdf

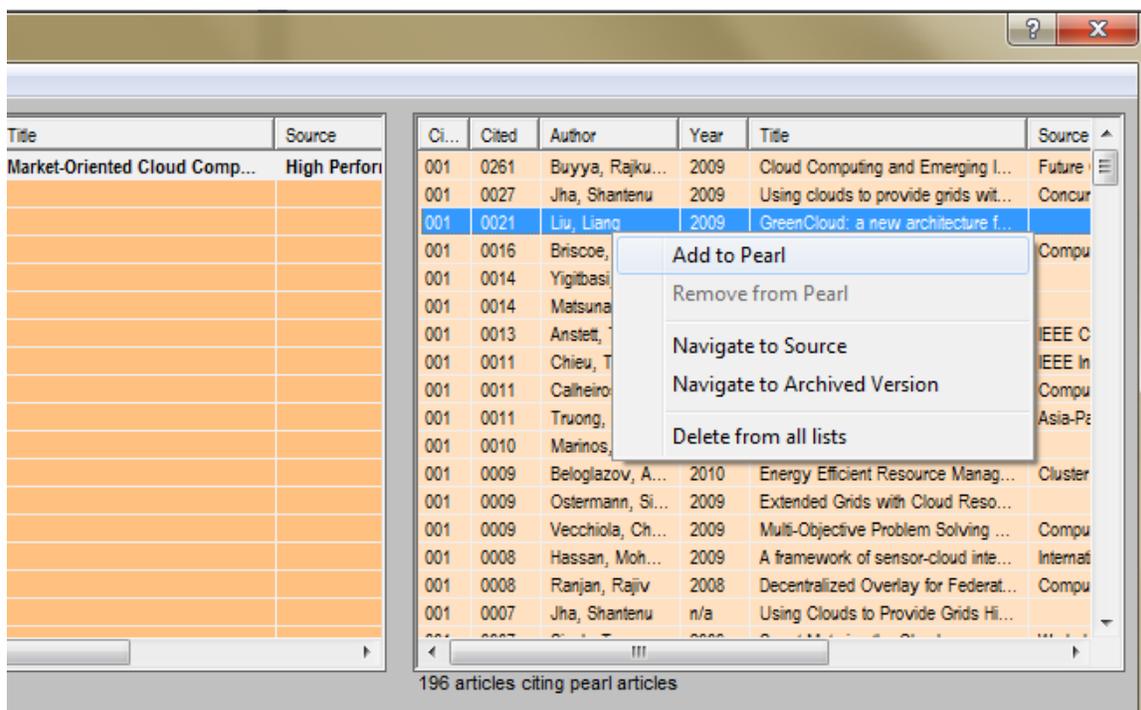
## Understanding Overlap Scores

Let's now return to the overlap scores shown in the first column of each list view. As more articles are added to the Pearl the overlap scores of each record will start to increase, with some increasing faster than others. The more pearl records an article relates to, the bigger its score. Hence, when a list is sorted in descending order by the overlap column, then articles that are cited by (left list) or cite (right list) the most pearl articles will appear both at the top of the list. We believe this makes sense, from a relevance point of view. For example, if you have identified 5 articles as relevant to your topic and another article cites all five of them, then it seems intuitive to expect that this article is also relevant. In contrast, if an article in the same scenario is cited just once by the Pearl articles, then it seems likely to be much less interesting.

## Adding further items to the Pearl

So, in order to take advantage of the overlap score we must add further items to the Pearl. This can be achieved in two ways. First, further items can be added from the Search results list, as described above. Note that you can re-define your query and retrieve new results as often as you like. The second method, which should become more important as the Pearl list grows in size, is to add items directly from the Cited or Citing lists. This is achieved in the same manner: by right-clicking on the record and selecting Add to Pearl from the menu. In the example below, we are about to add the article entitled "GreenCloud: a new architecture..." by Liu.

Each time you add a new Pearl article, using either method, the three lists will be updated. Any new cited articles will be added to the left list and any new citing articles added to the right list. If a cited/citing article is already a member of a list, then its overlap score will be incremented. This is important and is where the term *citation chain aggregation* comes from. You can imagine a citation chain as a graph where a single Pearl article is connected to its ancestors on one side (cited articles) and its descendents on the other (citing articles). What is represented in the three list view is the aggregation of the citation chains belonging to all of your Pearl articles. The overlap score is, in essence, the extent to which particular citing and cited articles repeatedly occur. Hence, the greater the overlap score, the more likely an article is to be relevant.



By gradually adding articles from the left or right lists, you are engaging in a form of citation cycling. For instance, if you know the Pearl article well, you might start by picking one or more important references from the Cited list. After each addition, you should check both peripheral lists to see which articles have high overlap scores. This backwards and forwards process is akin to a game of tennis – citation tennis! You can of course view and add citing and cited articles in any order you like.

The screenshot below shows the state of the session after four articles have been added. We can see that there are several articles at the top of each list that have an overlap score of two. Many of these are Buyya’s own articles, but we can see, highlighted in green in the right list, that there is a paper by another author, Ozer, that cites two of Buyya’s articles. These two articles are highlighted using bold in the Pearl list.

Bold font is used to show relationships between lists whenever you click on an article record. As we've seen, if you click on an item in the Citing or Cited list, then the related items in the Pearl list will turn bold. On the other hand, if you click on an item in the Pearl list, then related items in both of the peripheral lists will turn bold.

The screenshot shows the Oyster Search v1.00 beta interface. It features three main columns of article lists: 'Cited', 'Pearl', and 'Citing'. The 'Pearl' list is highlighted, and related items in the 'Cited' and 'Citing' lists are bolded. An 'Article Summary' window is open at the bottom, displaying the following information:

```

AUTHORS: Ozer, A.; Ozturan, C.,
YEAR: 2009
TITLE: An auction based mathematical model and heuristics for resource co-allocation problem in grids and clouds
SOURCE: International Conference on Software Computing, Computing with Words and Reasoning in System Analysis, Decision and
  
```

You may have noticed that some items in the left and right lists are now greyed-out. This means that they have already been added to the Pearl. This is a good place to point out what may seem obvious to you already – any article can appear in one, two or all of the three lists, depending upon its relationship to the Pearl. Another thing that might still seem confusing is what the overlap scores in the Pearl list mean. Well, this value represents the number of times an article is cited by other Pearl articles. In the example shown above, two articles by Buyya are each cited by two other Pearl members. Hence, in this context the overlap score gives a measure of relative importance between Pearl members.

In many cases it is possible to view the full text of the article, in addition to the basic meta-data. By right-clicking on the article record, two methods are provided for retrieving full text. The first is to Navigate to Source, which will attempt to access the article from the publisher's web-site. The second is to Navigate to Archived Version, which will access a copy stored elsewhere (e.g. an institutional archive or author's homepage). If either of these options fail, then all is not lost – CCA will open a browser tab using a Google Scholar query which will hopefully lead you to an alternative full-text source.

You can continue your CCA session for as long as you feel you need to - there are no strict limits on the number of records you can download. Note, however, that should you add a highly cited article (e.g. thousands of citations), then it may take a moment or two to retrieve all of these records from the MAS database. Whenever adding an article that has more than 500 citations, a warning is displayed giving you the option to limit the download to the 500 most highly cited citations. Bear in mind that limiting the retrieval of citing articles might prevent you from acquiring important information. If you can wait for the full download, this is always better. During a large retrieval, the caption on the

title bar of the Form will provide you with a dynamic indication of the time remaining until completion.

Finally, there will be times when you change your mind about an article. If you want to remove an article from the Pearl list, then right-click on its record and select **Remove From Pearl** from the menu. Note that this action will also remove any cited or citing articles that are exclusive to that article. The overlap scores of any other cited/citing articles will also be decremented by one.

At anytime during your session you may save your Pearl list for future reference, using the Save options in the Session menu. At the end of your session, you will also be prompted to save your final list. Files are saved in a well-known format known as **RIS Export**. This can be loaded into many existing reference managers, including EndNote and RefWorks. It can, of course, be re-loaded into CCA, should you wish to continue your search session.

### ***Import Bibliography (Hot Start)***

Up to this point, we have described the main features and usage of Oyster. However, the previous examples were predicated upon the assumption that you know very little about your search topic. The Import Bibliography start route assumes that you already have access to a list of articles that are relevant to your research, with the goal being to expand this list with new articles that are also relevant. As we discuss later, this might be reference list for a paper you are writing or a bibliography that you have been building for a new project.

As described, when you start Oyster using this option, you will be prompted to identify the source file. This must be saved in the well known **RIS export** format. Most popular reference management software (e.g. EndNote, RefWorks) can save bibliographies in this format and, as described above, it is the default format used by CCA when saving a Pearl list.

When you select a valid file, each record will be loaded into the Pearl list. All cited and citing articles will also be retrieved from the MAS database. This might take anything from a few seconds to several minutes, depending upon the number of articles and the number of citations that must be retrieved. Once loaded, you can continue to use the Oyster tool according to the procedures described earlier.

## **Summary and trouble shooting**

Please make sure you read the instructions in this document carefully before beginning to use the software. The Oyster tool has been tested thoroughly and should not malfunction under normal usage. However, no software is perfect and if you find any problems, persistent or otherwise, please do not hesitate to email us.

The most likely issues are likely to be related to the performance of the MAS web service. For instance, you may find performance slower than normal at certain times of the day. Whilst we have no control over this data service, please do let us know if data has been unavailable for an extended length of time (i.e. several days or longer).