



Software
Sustainability
Institute

SSI – JournalTOCs review

Project Title	SSI
Document Title	SSI – JournalTOCs review
Authorship	Mike Jackson
Document Filename	SSI-JournalTOCsReview.doc
Document Version	1.0
Distribution Classification	PUBLIC
Distribution List	<i>Public</i>
Approval List	<i>Project Management Board (SSI), Santy Chumbe (Heriot-Watt)</i>



Document History

Personnel	Date	Comment	Version
MJJ	16/06/10	Version 1.0 for public release.	1.0

1	Purpose of this Document	2
2	General comments	2
2.1	Layout and user classes	3
2.2	Service name	4
2.3	Navigation and hyperlinks	4
2.4	Browser resizing	4
2.5	Help	4
2.6	Fonts and text	5
3	User perspective	5
3.1	Registration	6
3.2	Browsing subjects and publications	6
3.3	Viewing the TOC page for a journal	7
3.4	Search	8
3.5	MyTOCs	8
3.6	Browse, search and MyTOCs window layout and interaction	9
3.6.1	Default	9
3.6.2	Browsing	10
3.6.3	Search	10
3.6.4	MyTOCs	11
3.7	Help	11
3.8	Quality of Service	12
4	Publisher perspective	12
4.1	What the site does	13
4.2	Copyright and licensing	13
4.3	What the site offers publishers	13
5	Developer perspective	14
5.1	What services are available	14
5.2	How to use the services	14
5.2.1	Case studies and examples	16
5.3	Quality of service	16
5.4	Copyright and licensing	17
6	Internal developer perspective	17
7	JournalTOCs and its competitors	18
7.1	Web of Science	20
7.2	TicTOCs	20
7.3	Scopus	21
7.4	Suncat	22
7.5	Zetoc	22
7.6	How JournalTOCs could compete	23
7.6.1	Usability	23
7.6.2	Promoting JournalTOCs	23

1 Purpose of this Document

A follow-up from the JISC RI Software Sustainability workshop in 12/05/10 was a request by Santy Chumbe (s.chumbe@hw.ac.uk) of Heriot-Watt University relating to the JournalTOCs WWW site (<http://www.journaltoocs.hw.ac.uk/>). This site provides an on-line web form for searching the tables of contents of over 14,000 online journal and other publication services, these being provided via publishers' RSS feeds. It also provides services – accessed via a RESTful API – that allow searches over these tables of contents. Santy's concerns were the usability of the site, how to more effectively compete against commercial offerings and documenting the services the site offers for developers.

This document provides a review of JournalTOCs undertaken in WB24/06/2010. It is done from the perspective of:

- Users who use the on-line web form to run searches across the registered journals.
- Publishers who wish to contribute to the site.
- Developers who wish to call the services from their own applications.
- Developers who wish to understand how the services work and how they could extend or change them if they had access to the source code e.g. developers joining the JournalTOCs project.

There are also general comments on the web site that relate to all classes of users and comments on how JournalTOCs might compete against other offerings.

To summarise, no fundamental problems were encountered and most of these can be addressed by redesign of the web site and its interfaces. The main recommendations are:

- Restructure the web site pages to recognise the different informational needs of the various classes of user – online user, publisher, developer.
- Restructure the web site pages to remove the three-column format and so allow the display to be less cluttered and the search results to be more easily navigable.
- Attempt to clarify the policies regarding usage of the information (e.g. copyright and licencing) exposed via JournalTOCs services and publish this on the web site. This information is important for developers who will develop services based on those of JournalTOCs.
- Emphasise successes on the web site and strengths compared to competitors'.

The following sections contain the review in detail. Please note, I've included every single issue, no matter how trivial, that I've encountered, though I do highlight the most important.

Please contact me at m.jackson@software.ac.uk if you have any queries.

2 General comments

Looking at the site generally, there are two main issues.

-
- The use of fonts and styles for emphasis and hyperlinks is inconsistent (e.g. bold, underlined, capitals are each used for hyperlinks, bold is also used for emphasis).
 - The site interleaves marketing, user, developer and publisher information in a haphazard way. Each of these groups have specific information requirements and the site should be structured with these in mind so they can easily access the information they need.

A detailed overview of issues encountered is now given with suggestions. Most of these apply to the interface. A few though apply to the underlying services.

2.1 Layout and user classes

The web site is a mix of marketing info (what the service offers and its unique features), information for users of the online browse/search/MyTOCs functionality, information for publishers, information for developers wishing to use the service and some information on how it works. This information is intermixed in the home, about and help pages. For example, the help page mixes information on:

- How to run searches and save favourites – for users.
- Link to API – for developers.
- Contact info – for all.
- What it does, what it offers users, developers and publishers – for all.

Such intermixing can make it difficult for specific user classes to find just the information for them without distractions – each of these user classes have different information requirements. I'd propose a restructuring that addresses the various classes of user, for example, redesigning the main pages to have something like:

- Home – the home page. Very high-level welcome information promoting the number of articles and journals available and some of the available publishers. Certainly this text from the about page – “JournalTOCs is the largest, free and searchable collection of scholarly journal Tables of Contents (TOCs) in the world. It contains TOCs for over 14,000 journals collected from over 500 publishers.” should be on the home page as it's a great claim!
- About – what JournalTOCs is and the number of journals, publishers involved, other services built on this, like WattJournals etc, sub-sections with very high-level information (nothing on how to use it) for users, developers and publishers, with links to blog and JISC. Basically, selected content from the home, about and help pages restructured, cut down and rewritten.
- MyTOCS – info about MyTOCs and sign in/register form.
- Developers – detailed information on services available, how to use these, quality of service, copyright and licensing, case studies. Basically, content from the API page with additions and restructuring.
- Publishers – detailed information for publishers on what the site does with the information from their RSS feeds, how it allows third-parties to build applications using this and a contact in case they wish to discuss copyright or licencing. Basically, content from the help and about pages with additions and restructuring.
- Help – help for users.
- Contact us – info on e-mails (who developers, users and publishers should contact and about what issues) and a blog link.

The search box should be permanently visible – see 3.6.

2.2 Service name

The site header bar says JOURNAL TOCs, the home page says JournalTOCs, the about page says JournalTOCs, the help page says Journal TOC and journalTOCs. It can help with service identity to adopt a consistent naming throughout – this includes which letters are capitalised and which are not and where spaces occur e.g. at EPCC we were involved in a project called BEinGRID funded by the EU and the project guidelines stated it was always to be written as BEinGRID, never Beingrid, BEinGrid or other variants.

2.3 Navigation and hyperlinks

The use of fonts and styles is inconsistent. Sometimes bold is used for hyperlinks, sometimes for emphasis. Sometimes underlining is used for hyperlinks. Sometimes no special markup is used and whether something is a link depends on the context of the text (e.g. in the publishers/subjects table or in the search results list). This means the user may click words that aren't meant to be clicked (and so see nothing happen) or not click on words they are meant to click (and so miss results or information they need). A consistent link format should be adopted throughout. Changing text to black and displaying links in blue (as done in Wikipedia or SourceForge) would be a suitable format.

Sometimes a phrase takes the user to a place they might not expect e.g. “API” on the help page pops up a new browser with the Wikipedia entry for API when they might expect to be taken to the JournalTOCs API page. Or, certain links cause unexpected behaviours e.g. the “OPML” link on the help page pops up a small window about OPML yet the other technology links on that page (“API”, “RSS”, “XML”) open the associated Wikipedia page in the browser. I'd recommend that the “OPML” link open the associated Wikipedia page also (info on using OPML can be added to the help for users – see 3.7)

2.4 Browser resizing

If the browser window is made smaller then text in the middle panel can be truncated. This happens to the URLs on the API page, for example.

2.5 Help

It is both helpful and consistent if all widgets (buttons, fields, checkboxes, radiobuttons etc.) have a tool-tip.

At present, questions, registration requests, feedback and files from publishers are requested to be sent to S.Chumbe@hw.ac.uk. Bug reports are directed to santiago@icbl.hw.ac.uk. It can help separate the project from any individual if a project e-mail is set up e.g. journaltocs@hw.ac.uk, or, having a project-specific domain name, support@journaltocs.ac.uk.

2.6 Fonts and text

It can be clearer if initial capitals are only used for proper nouns e.g. names of services or products.

On the registration confirmed page, change “free journal tables of contents Service” to “free journal tables of contents service”

On the Search selection options, change, “for Journals by Title/ISSN” to “for journals by title/ISSN” and “for Articles by Keywords” to “for articles by keywords”

Headings should be bigger or more distinct than their accompanying text.

There are a number of text colours used. Besides from deep blue, many of the pages use orange headings or deep orange fonts (e.g. JournalTOCS on the About page) and green text is used on the page displayed after a user has registered for the first time or on the registration confirmed page. It's better to be consistent in the use of colour (e.g. make the green text blue also) unless there is some specific intent. Also, a high contrast between text and background is preferable, and gives less problems for partially-sighted users, so using headings the same colour as the text (e.g. deep blue) rather than orange would help.

3 User perspective

This review is from the perspective of a user wanting to run on-line searches for journals or articles of interest and use the MyTOCs feature which allows users to register and save links to their favourite journals.

As a user I would want to know:

- What the site offers me.
- How to use it.

The home page makes it clear what the site offers – the ability to search and browse journal tables of contents and abstracts – and the about page provides further information on this.

The home page and help page make it clear how to use the site – though the information is spread across these two pages and some information was missing (see 3.7 for example). However, it was generally easy to use without reading the help page. Some features required clicking to find out what would happen and in some cases it wasn't clear by looking at the interface what exactly I, as a user, could or could not do. As a software developer I'm used to just clicking on things and looking for hyperlinks even if they're not clearly marked. A less computer-literate user might be more cautious or not even think of doing that. The user is well protected from the internals of the service. No matter what I clicked or entered I didn't see any cryptic error messages anything relating to how the service works. The main issue is the layout of the interface and how, when browsing through the information available – from subject and publisher lists to lists of journals and articles to individual tables of contents – things didn't appear where one might expect.

A detailed overview of issues encountered is now given with suggestions. Most of these apply to the interface. A few apply to the underlying services.

3.1 Registration

It can aid navigation if users always know where they are. The user registration page and the registration confirmed page would help this if they had clear headings e.g. “Register”, “Registration confirmed”.

The registration and registration confirmed pages are in a single column layout yet the sign in page is in the three column layout. A single column layout for the sign in page too would make the login-related pages all consistent. Consistency can assist user understanding as to what is related.

I would prefer it if the login failed message to appear under the sign in button.

It can be easier for users to understand which actions complement each other if common opposites are used e.g. add/delete, show/hide. For JournalTOCs, adopt “Sign in/Sign out” or “Log in/Log out” instead of “Sign in/Log out”.

If I click on the Sign In button without entering a username or password I’m taken back to the home page. I’d expect a login failed message.


On the Registration Confirmed page clicking on the Back to JournalTOCs Home Page link gives a “page not found” error. The broken link is <http://www.journaltoCs.hw.ac.uk/www.journaltoCs.hw.ac.uk>.

3.2 Browsing subjects and publications

It’s common in tables whose column headings have hyperlinks that clicking on the column heading will:

- Either display more information about the data in that column.
- Or, reorder the rows of the table sorted according to the order of data in that column e.g. alphabetical order. Typically if this functionality is supported then clicking the table column again can be assumed to sort the column in reverse order.

The Subjects table is inconsistent with this which may cause the user some initial confusion as things don’t behave the way they expect. If the user clicks on the Subject heading then the table is sorted by the number of journals from greatest to least. If the user clicks on Subject again it’s sorted by subject in alphabetical order. It would be better if both the Subject and Journal TOCs column headers were clickable and clicking on Subject were to sort by alphabetical order and then reverse alphabetical order if clicked again, and clicking on Journal TOCs were to sort by number of journals, greatest to least, then least to greatest if clicked again.

It would make it more clear to the user that the columns are sortable if, instead of having the users click on the column names, a small up arrow/down arrow icon () were presented in each column

The JournalTOCs column should be renamed “Number of journals” since it’s clearer what that means.

The same comments apply for the Publishers table.

The sentence “Browse Journal TOCs by Publishers or by Subjects” doesn’t make it entirely clear that one can see either Publishers or Subjects tables. It might be better to have separate buttons to select these.

The display doesn’t make it clear e.g. by underlining, that the subjects and publishers in the table and the names of journals are hyperlinks and are clickable. It does change mouse pointer when the user moves the mouse over it but the user has to think of doing that first.


After clicking on a subject or publisher, the list of matches appears to the left. This is counter-intuitive since usually when hierarchically browsing information the it appears on the right (think of the file browser for example). It would be preferable just to open a new page with the list. This relates to the suggestions of 3.6.

It might be useful to allow the user to see a list of all publishers or all journals.

3.3 Viewing the TOC page for a journal

The TOC page for an individual journal is a bit confusing. The journal name and information is lost between the MyTOC checkbox and the GetTOC feed link, the horizontal line and the orange Table of Contents heading (the font of which is larger than the journal title). It would be better if the journal title is printed first and with the largest font – that is the information that provides the context for the user, then the checkbox and feed link as smaller widgets and the Table of Contents as a heading with a smaller font.

If possible, it would be useful to have info as to whether full article can be downloaded for free or requires subscription or registration with the site that publishes the article. It could be wasteful of the users’ time to go to the publishers web site only to then discover they’d have to register.

The display doesn’t make it clear e.g. by underlining or an external link icon (e.g.  as available at <http://upload.wikimedia.org/wikipedia/commons/2/25/External.svg>), that the journal name and the abstract names are hyperlinks and will take the user to the journal pages. It does change mouse pointer when the user moves the mouse over it but the user has to think of doing that. The external link icon may be preferable since it makes it clear that the user will be leaving the JournalTOCs site.

Nothing on the TOC page or in the help page explains what RefWorks is or does.

By chance I navigated to ApJ Letters Latest Papers – ISSN (Print) 1538-4357. The comment says that “There are no recent issues published by this journal in the last six months.” But their web site – <http://iopscience.iop.org/2041-8205> - does cite more up-to-date publications. Not sure what the cause of this is but it is important for users to know the distinction between a journal that has no up to date publications and one that does have up to date publications but which aren’t available via journal TOCs. Perhaps it might be more accurate for the site to say “No recent issues available from our service”.

For browsing an individual abstract, after clicking on a search result, it would be more consistent to adopt a similar format to the table of contents view i.e. journal title, then the abstract, then the “View full table of contents” link. This link should also be made more prominent – or made a button.

The order of abstracts is different from that seen if clicking on the Get TOC Feed link to see the feed directly from the publisher. For articles that have the same publication date/time the order is reversed. This could be confusing for users if they browse from the abstracts list to the raw feed.

3.4 Search

I tried very generic search terms which have the potential to throw up a large number of hits e.g. “the”, “*” and “c*” each as searchers. The page refreshed but that was all. There were no search results nor was there any indication of 0 hits returned. If the service has a maximum number of entries it can return (e.g. embedded in the RSS examples in the technical doc there’s an XML comment “(maximum number of returned articles is 2000 (300 if your query includes wildcard character *)” or ignores certain search terms then these should be documented.

All the ways to control the search should be documented for the user e.g. use of OR, which is only documented in the default feed page for the base URL (<http://www.journaltoocs.hw.ac.uk/api/>).

Phrase searches didn’t behave as expected. If I search for journals with the title phrase “Advanced” I get 0 hits. If I search for “Advances in” I get 162 hits. I’d expect one word phrases to be treated like keywords.

The search seems case-insensitive. This should be documented.

It would be useful if the user could specify a search just over the journals in their MyTOC.

3.5 MyTOCs

Sometimes MyTOCs would display an OPMLExpert icon and an “i” icon. At other times it would display a “Expert “MyTOCs” in OPML format” link and a “?” button. The switch of formats could prove confusing for users.

The “i” icon is a standard way of presenting help – see below – and should be kept but I think it would be more consistent to just have a “Download TOCs” button rather than an image icon. I initially considered “Save” or “Save As” but there’s a risk that user’s might think that just saves them onto the server.

Some links have a star icon with a tool-tip “Mark as read” I clicked on these and they disappeared. Nowhere does it explain what this does or is. After looking at ticTOCs I assume that this is journals that have had updates since my last login.

The OPML pop-up has a section “How does it work” but actually it describes “How to use it”

The OPML pop-up says:

- 1.Login to Google Reader
- 2.Click Manage Subscriptions
- 3.Click the Import/Export tab
- 4.Browse for your OPML file
- 5.Click Open
- 6.Click Upload

I tried this but Google Reader didn't have Manage Subscriptions at the top level. I had to click the V icon by the Subscriptions header to be taken to that page. And, instead of browsing for my OPML file and clicking Open – there was no Open button – I had to:

“Click Browse... and select your OPML file.”

“Click Upload”

The checkbox isn't the best widget to use for MyTOC entries as its behaviour is inconsistent with what users might expect. When a user clicks a checkbox they expect it to remain but the check mark to appear or disappear. If they click the checkbox here the entry disappears. It may be better to use a button/icon by each entry e.g. a (X) or trash-can icon. Alternatively, see 3.6.

3.6 Browse, search and MyTOCs window layout and interaction

The layout of the windows in respect to browsing, searching and MyTOCs could be improved. For example, JournalTOCs offers the ability to browse or search information that is hierarchically structured from publishers/subjects to lists of matching journals or abstracts to individual tables of contents. In JournalTOCs the general info (publishers/subjects) and the specific information (contents) is presented in the centre and the intermediate level information is presented on the left. This can make it confusing to users who are unclear as to where things might appear and how they relate. The three column format also means the display can seem rather cramped.

3.6.1 Default

I'd propose adopting a vertical rather than horizontal layout. The upper panel contains Subjects, Publishers and Search buttons, along with the various buttons to control the search. This should be visible on the home page.

On the home page the lower panel can contain general information about JournalTOCs, much like the central panel does at present.

Subjects Publishers Search ...
JournalTOCs is...

3.6.2 Browsing

For browsing, when the user clicks Subjects or Publishers, the lower panel can display the subjects / publishers tables.

Subjects Publishers Search ...
<p>Subjects (133 total)</p> <p>Abstracting and Indexing</p> <p>Advertising and Public Relations Journals</p> <p>...</p>

The table could either be paginated, say 20 rows at a time, with Next and Previous buttons to navigate back and forth, or just one big table, with a scrollbar if needed. In either case, it can be useful to see the total number of results. Ideally, the user would be able to switch between these two views. It would also be useful if the user could sort the results by title/name.

When the user clicks on a publisher / subject row the entries for just that publisher / subject can be displayed in the lower panel e.g.:

Subjects Publishers Search ...
<p>Subject – Advertising and Public Relations Journals (13 total)</p> <p><input type="radio"/> International Journal of Advertising</p> <p><input checked="" type="checkbox"/> International Journal of Market Research</p> <p>...</p>

The same comments apply as to the presentation of the table as for the subjects / publications. As for the results lists at present, the table could provide a column, MyTOCs, with a checkbox to indicate whether it's added or not – this is not done on the current interface and it was unclear when I first started using JournalTOCs what the checkboxes by the search results or lists of journals by subject did. Only by clicking did I see that they added/removed them to/from MyTOCs. Select All, Clear All, Add To MyTOCs and Delete From MyTOCs buttons would be useful to allow the user to do batch additions and removals.

3.6.3 Search

On running a search the results can be displayed beneath e.g.

Subjects Publishers Search ...
<p>Search results for “genetic” (565 total)</p> <p>1. ...</p> <p>2. ...</p> <p>3. ...</p> <p>...</p>

Similar comments apply to the presentation of the table and MyTOCS-related buttons as already described.

Clicking on a search result could take the user to a new page where the publication’s table of contents, or the individual abstract are presented.

3.6.4 MyTOCs

I don’t think it’s necessary to have the MyTOCs permanently on display since that, to me, seems a separate feature. Besides, the interface is good at making it clear what journals are members of MyTOCs. I think this should be accessed via a link on the menu bar.

Subjects Publishers Search ...
<p>MyTOCs (12 total)</p> <p>x International Journal of Market Research</p> <p>...</p>

Similar comments apply to the presentation of the table and MyTOCS-related buttons as already described.

3.7 Help

There is no info on what RefWorks is.

There is no explicit statement that says what MyTOCs is. The front page implies it – “Select a TOC and save it in your MyTOCs folder” but it doesn’t say anywhere that it’s essentially the user’s bookmarked journals that are saved on the server.

There is no info on what the “Mark as read” star icon in MyTOCs means or what happens when it’s clicked.

The help says that when searching to “Please check the "in TOC articles" option”. The actual checkbox is called “for Articles by Keywords”.

The help page is a mixture of how to use the service, who to contact and what RSS feeds are. The RSS feeds info isn't really required to use the service via the web interface, it's more of interest to developers building on it or people wanting to understand how it works, so can be moved.

How help information is presented is inconsistent and distributed. There is some on the home page, more here, and some in the OPML pop-up help. This can make it difficult for users to find help and they may waste time searching for it.

I'd suggest rewriting the help page into the following sections:

- Intro – Number and types of resources available, general context information for users.
- Browsing – Info on how to browse using the Subjects/Publications tables and the format of the TOC page for an individual journal. What RefWorks is and does.
- Searching – Info on how to search.
- MyTOCs – What MyTOCs is and how to use it. What the “Mark as read” icon does. Info on how to save as OPML and view it in Google Reader.

Each heading could have a relative “id=#Tag” anchor defined. Then an “i” icon can be added alongside the Go search button, the Download MyTOCs button (3.5), and elsewhere. These could then kick-up a browser window scrolled to the relative link. This means that all the help is on one page, the way of accessing help is consistent and there is more context-specific help.

3.8 Quality of Service

It would be useful for both users and publishers to know what your policy is about keeping up to date e.g. what to do if a publisher shuts down a feed or renames it. Do you regularly check the feeds to see if they're still live? It would be useful for users, developers and yourself, to know this information.

4 Publisher perspective

JournalTOCs aggregates RSS feeds from publishers. These are aggregated by the JournalTOCs team. There is no requirement upon publishers to request they be included or to take any action. There is not the resource to include every journal available so the team assesses, based on their own awareness of the significance of journals, which to include. A new batch of journals, including 300 Open Access journals, are currently under assessment.

Though publishers do not contribute to or subscribe to JournalTOCs it is nevertheless important that certain information be provided for publishers, since, after all, it is their information that is being aggregated and re-distributed. Therefore, as a publisher I'd want to know:

- What the site does with the information from their RSS feeds.
- A contact in case they wish to discuss copyright and licencing.

- What the site offers publishers.

4.1 What the site does

The home page and about page provides a good overview of what the site does.

It's important that publishers be made aware that the information from their feeds is being further published by your site, via your APIs for developers, which relates to copyright and licencing.

4.2 Copyright and licencing

Nowhere is the issue of copyright or licencing addressed. Publishers' copyrighted material is being exposed on the site. One issue of importance is whether publishers explicitly know that third-parties can build their own applications using the service API and so use their contents and abstracts. They may allow fair-use and replication of their abstracts, they still retain the copyright after all, but may not be so happy if someone were to set up a commercial service based on their abstracts (e.g. commercial text mining, searching, analysis or other services). RSS feeds are new and this is a murky area (see, for example <http://www.feedforall.com/rss-copyright-debate.htm>).

I think it's best to be honest and state that this is a complex area, with links to discussions such as the above but also to cite that you have had informal discussions with some publishers and that the community is waiting for more clarification. Furthermore, it would be useful, and promote goodwill to both publishers and developers, if you were to state that you have been, and are continuing to, work with your University Technology and Research Transfer Office to ensure that JournalTOCs and any potential spin-offs do not infringe the rights of publishers. Likewise, you should say how any commercialisation by yourselves will only be done when you have devised a satisfactory resolution.

It would be a further symbol of goodwill if you were to have a statement that invites any publishers who have concerns they wish to discuss to contact you or, if they wish their feed to be removed from your service, that they are free to request this.

A statement on these issues to make any assumptions explicit would be useful for both publishers and developers.

Discussions with ATIRA about PURE that have been mentioned would benefit from such clarification.

4.3 What the site offers publishers

Nowhere does it say what JournalTOCs' usage of, and exposure of, my RSS feeds would give me as a publisher. It may help to offset any concerns a publisher has, in terms of licencing and copyright, by explaining that the service you offer may increase their readership, their subscriptions, and possibly their profits.

In return, seeking endorsements from publishers whose feeds you already aggregate may help with marketing.

5 Developer perspective

As a developer building applications using the JournalTOCs services I want to know:

- What services are available.
- How to use these.
- What is the quality of service.
- Copyright and licensing conditions.

5.1 What services are available

The API page provides information as to what services are available, with examples.

5.2 How to use the services

An overview is described on the API page with examples – this is enough to get a developer started. There is a comprehensive reference guide to the API in the technical docs. This separation is a good approach, the developer can get things going quickly without being overwhelmed by every available option.

It would be useful to have a link to the root of the technical doc from the API page.

The information on the API page could be better structured. At present there are the three service URLs, then further down the page the examples. Grouping these together and using task-oriented headings would be clearer. For example:

JournalTOCs for Developers

...

Link to technical doc

...

Search for journal titles

URL: <http://www.journaltoocs.hw.ac.uk/api/journals>

Use this call to get the TOC RSS feeds of a specific journal or to get all the journals whose titles match your query string. The results are returned in RSS 1.0 format.

Examples:

- <http://www.journaltocs.hw.ac.uk/api/journals/library> gives a list of journals whose title contains the word library.

Find articles matching a query string

...

Get a list of your journals

...

Get a list of custom search results

...

Features common to all services

...

Clicking on the default service URLs takes the user to default feed pages which explain what the service does and then tells the user to “Go to <http://www.journaltocs.hw.ac.uk/index.php?action=api> for further information.” It would be useful to make this a hyperlink. It would also be useful to link to the appropriate page in the technical doc.

The examples on these default feed pages are different from those on the API page and these are different from those in the technical doc. It could be clearer for the user if these were made consistent (so they’d recognise the examples across the three docs).

The articles example - <http://www.journaltocs.hw.ac.uk/api/articles/corrosion> - on the API page fails with an error:

```
Fatal error: Allowed memory size of 16777216 bytes exhausted (tried to allocate 1650656 bytes) in /var/www/html/icbl/journaltocs/api/articles_index.php on line 211
```

Relating to documenting the ways the search can be controlled in 3.4, all the ways to control the search via URL parameters should be documented for the developer. At present this is spread across the help page, the API page, the technical doc and the default feed page for the base URL (<http://www.journaltocs.hw.ac.uk/api/>).

The notes section is a bit small and has no clear heading as such. I suggest making it a heading. You might rename it to “Features Common to All Services” (or another similar title). In this section it would be good to have a link to the Wikipedia RSS page and a direct link to the RSS 1.0 specification (<http://web.resource.org/rss/1.0/>) since that is the output format used by the services. This would save having the text “The results are returned in RSS 1.0 format” for every service and so cuts down the verbosity.

The RSS elements used by the API for search results – available on the blog at <http://www.journaltocs.hw.ac.uk/API/blog/?p=90> – would be more appropriate in the technical doc (<http://www.journaltocs.hw.ac.uk/docs/index.php>) as this keeps all the service-related reference doc in one place. I don’t think blogs are suitable for persisting technical docs – to me blogs are more transient, conversational resources.

Relating to the use of very general search terms in 3.4, running http://www.journaltocs.hw.ac.uk/api/journals/* gives a page with 0 results but with the comment “0 Hits! Sorry, found nothing for your query: * (do not use common words in your query, such as "journal")”. This implies some query filtering is going on to strip out very common search terms. If so then this should be documented. As a developer it’s important to know what searches are run but return 0 hits because there are no matches and which are not run because the search terms are too general.

The example RSS responses from example queries are very useful for developers.

The institution examples on the API page are inconsistent with the examples on the API page and in the default feed page at <http://www.journaltocs.hw.ac.uk/api/>. It states the URL is <http://www.journaltocs.hw.ac.uk/api/institution/yourServiceID>, but the examples are similar <http://www.journaltocs.hw.ac.uk/api/HeriotWatt/corrosion%20steels>, which is confusing. The examples should be consistent with the general form and cite “institution” e.g. <http://www.journaltocs.hw.ac.uk/api/institution/HeriotWatt/corrosion%20steels>. This also means they’re consistent with the other services e.g. .../api/call/arguments.

The API page should be consistent with the technical doc and state that the institution-specific search is only available using “HeriotWatt” or “TechXtra” or “ir_bgs” for now.

The <http://www.journaltocs.hw.ac.uk/api/institution/HeriotWatt/%22bayesian%20filtering%22> example on the institution technical doc page fails with an “Internet Explorer cannot display this feed”.

5.2.1 Case studies and examples

It can be useful for developers to know what others have done using JournalTOCs. This can also be useful for marketing. The case studies forwarded by e-mail (e.g. WattJournals - <http://www.hw.ac.uk/library/wattjournals.php> or TechJournalContents - <http://www.techxtra.ac.uk/techtocs/>) provide good examples of this. Developers would also, of course, benefit from knowing how these services were built.

5.3 Quality of service

A developer building an application based on JournalTOCs may want to know what quality of service they can expect as this will affect their application (especially if they are offering their application as a service and wish to make quality of service statements of their own).

For example, the API page only cites that 1000 journals can be saved in a MyTOCs folder. Also, I ran a search using the search form for all articles citing “data” which returned 39378 hits. Running the same query via the journals service - <http://www.journaltocs.hw.ac.uk/api/journals/data?output=articles> - gave an error:

```
Fatal error: Allowed memory size of 16777216 bytes exhausted (tried to allocate 120686 bytes) in /var/www/html/icbl/journaltocs/api/lib/api_fns.php on line 551
```

Embedded in the RSS examples in the technical doc there's an XML comment "(maximum number of returned articles is 2000 (300 if your query includes wildcard character *". This should be documented.

But, is there an upper bound on the number of searches that can be run in a day? What is the lifespan of the service, how long can it be expected to remain operational? Is there a means by which developers can be notified (e.g. via e-mail) of scheduled downtimes etc.

It is OK to have a very low quality of service statement (e.g. "no guarantees as to lifetime") so long as it's accurate, but such a statement should be available.

The blog contains some information that might be of use to a developer (e.g. <http://www.journaltocs.hw.ac.uk/API/blog/?p=143> cites issues when getting 12,000 search results and in handling HTTP response status codes). This would better be published as part of the developer docs so they're aware of the current state of the service and any limitations.

5.4 Copyright and licensing

Developers can access the licence of the API via the API page. It would be better if there was a heading drawing attention to this.

The comments for publishers in 4.2 are relevant to developers too. There is no statement about the conditions affecting use of the tables or contents or abstracts that the service makes available. Can these be manipulated or altered? Do the original authors or publishers have to be credited? Can it be used for commercial gain?

It is important for developers to be aware of any conditions in the use of JournalTOCs or the data it provides.

6 Internal developer perspective

As a developer wanting to understand how JournalTOCs works and its internals I want to know:

- How JournalTOCs works.
- What are its main components.
- Apart from RSS, HTTP and OPML, what languages, technologies and products are used to build the service endpoints, manage the feeds from the journals, manage MyTOCs, build the web site.
- How it processes feeds from each journal.
- How is journal info, feed info and MyTOCs stored on the server.
- How it runs its searches.

There is no such information available on the web site. A browse through the blog might give up some information (e.g. the fact that Xerces is used) but apart from that there's nothing.

7 JournalTOCs and its competitors

Santy forwarded a list of competitors and their perceived disadvantages. Rearranging these as a feature grid is a common way of highlighting how a service compares to its competitors (and can appear less biased than just a list of why one service is good and others not). This is an example, augmented with more information from the products' web sites.

	JournalTOCs	Web of Science	JournalTOCs	Scopus	SUNCAT	ZETOC
Cost	Free	Subscription from UK HE/FE institutions, research councils and NHS. From £1750 to £18K depending on organisation.	Free	Annual subscription – depends on institution size. Relatively expensive	Free	Free to UK HE/FE institutions, research councils and NHS.
Update policy	Within a few hours. Depends on publishers feeds.	Up to 1 month lag.	Within a few hours. Depends on publishers feeds.	Daily.	Monthly-quarterly, less frequent – depending on contributing libraries.	Daily.
Longevity			Project ended 31/3/09. MIMAS will run service for 3 years. The University of Liverpool Library committed to maintaining data.			
Help and support	Online doc E-mail	Online doc E-mail E-mail list Telephone	Online doc E-mail	Online doc Training days E-mail Telephone	Online doc	Online doc E-mail Telephone
Number of publishers	556	Thomson Reuter's products	350	5,000 plus 318,000,000 web pages plus		

				5 patent offices		
Number of journals	14,111	10,000 journals 110,000 proceedings	11,000	18,000 (journals, publications, books, proceedings)		20,000 journals 16,000 proceedings
Number of records	2,000,000			41,000,000 (1,700,000 added every year) from 1823-2010. 23,000,000 patent records		37,000,000
Subject areas	133	256		4 (Life, Physical, Health, Arts and Social); 27 major areas; 300+ minor areas		
Information directly available	Journal titles, publishers etc, tables of contents.	Journal titles, publishers, tables of contents etc.	Journal titles, publishers, tables of contents etc.	Journal titles, publishers etc, tables of contents, full text.	Journal titles, publishers etc, not contents.	Journal titles, article titles, publishers etc.
Full text article access	Yes to open articles and those to which institution has access.	Yes to open articles and those to which institution has access.	Yes to open articles and those to which institution has access.	Yes to open articles and all those provided by Scopus.	No	No.
Search over	Journal title, ISSN, article title, abstract text.	Journal title, ISSN, article title, author, abstract text,...	Journal title, subject, publisher	Journal title, ISSN, article title, author, abstract text,...	Journal title, subject, ISSN.	Journal title, author, ISSN, publication year
Search types	Keyword Wildcard Phrase AND, OR	Keyword Wildcard Phrase AND, OR, NOT Constrain to subject area, date range, recency.	Keyword StartsWith	Keyword Wildcard Phrase AND, OR, NOT Constrain to subject area, date range, recency.	Keyword Phrase Boolean AND, OR Constrain to location, library, language, year.	Keyword Phrase
Search history	No	Yes	No	Yes	Yes	Yes

Save search results	Yes	Yes	Yes	Yes	Yes	Yes
APIs	RESTful endpoints for searches, browse user's bookmarked contents, browse institution's	OpenURL endpoints – access from partner sites	Planned	OpenURL endpoints – access from partner sites Search API – basic doc and citation searches for registered non-subscribers. More comprehensive functions for subscribers.	None	Z39.50 inter-database connection protocol OpenURL RESTful endpoints SOAP-based web services. DNS and IP checks are done.
Alerts	No	Yes	Planned	Daily, weekly, monthly.	No	Daily

7.1 Web of Science

<http://wok.mimas.ac.uk/>

MIMAS at The University of Manchester, JISC, British Library.

ISI Web of Knowledge Service for UK Education is a service based on Web of Science by Thomson Reuters (http://thomsonreuters.com/products_services/science/science_products/a-z/web_of_science). Web of Science is a comprehensive publication search and citation tool for researchers. Articles can be browsed and searched. Authors and citations are cross-referenced.

There are simple and advanced search forms. The simple form allows searches to be expressed in terms of keywords and boolean operators. The advanced form allows specification of a search phrase with boolean operators and other search control commands. Searches can specify timespans, citation indices to search over, language and document type.

Search results can be refined by category, subject, document type, year etc. The title, author and abstract etc can be viewed for each result and these can be printed or saved.

The search and results windows are more aesthetically pleasing than JournalTOCs mainly due to the lack of a multi-column format – a tabbed panel is used instead – and the search results span the full width of the display. The display seems less cramped.

7.2 TicTOCs

<http://www.tictocs.ac.uk/>

Consortium including University of Liverpool, MIMAS, Cranfield University, Heriot-Watt University, JISC and others.

This site is in look and feel almost identical to JournalTOCs except for the colour scheme. It is slower for searches, however. Some of the recommendations I've made in the preceding text are already supported e.g. a page explicitly for publishers or the correct use of checkboxes in the MyTOCs list. Some other, more usable, aspects are:

- Ability to register and sign in without need for clicking on a link in a confirmation e-mail (though it's appreciated than confirmation links can beat spammers).
- Ability to select specific entries in MyTOCs, or select all, and remove or export these.
- Ability to select specific journals in search results or select/deselect all and add these to MyTOCs.
- Search results are presented as a scrollable list rather than pages of 20 entries at a time – giving the user the choice of either would be the most flexible.
- MyTOCs entries that have been updated since the user's last login are highlighted (I assume this is what the orange "Mark as Read" star icon does in JournalTOCs).
- Provides information on importing feeds into Outlook, Bloglines, iGoogle, myYahoo etc.

There is a plan to provide service APIs and use an API key to monitor service usage. As a developer I'm in two minds about this. On the one hand it can prevent abuse, on the other it can be problematic and affect portability if it's linked to a specific server (GoogleMaps API keys suffer from this problem).

7.3 Scopus

<http://info.scopus.com/>

Elsevier B.V.

This is a comprehensive publication search and citation tool for researchers, librarians, editors and publishers. Articles can be browsed and searched. Authors and citations are cross-referenced. This allows searches for which papers are cited by which others and the citations of a specific author's publications, for example.

The demo shows an on-line GUI (constructed using some toolkit like FLEX or Google Web Toolkit perhaps). Tab panels are used to group together information, widgets have rounded corners and the effect is more pleasing than JournalTOCs, though occasionally the interface can seem rather crowded with information.

Search can be free text across any field or can be composed by selection of boolean operators (AND/OR/NOT) and specific meta-data (e.g. Title, Abstract, Author, ISSN,...). Searches can be constrained to specific subject areas, date ranges or recency of publication. I'm not sure if users prefer searches specified in a form using drop down lists or the generic text field Google-style approach. It would be worth searching to see if any usability evaluation has been done on this.

The site lists benefits to publishers which includes increasing visibility, accessing a global audience, finding peer review experts, tracking performance, monitoring competitors, driving usage to their publications.

7.4 Suncat

<http://www.suncat.ac.uk/>.

EDINA, The University of Edinburgh, JISC.

A serials catalog allowing location of serials information held in 73 libraries across the UK. It is designed for researchers and librarians. The aim is different from that of JournalTOCs or Scopus for example in that it allows researchers to search serials to identify which libraries provide access to these.

There are basic and advanced search forms allowing search for subjects, titles, locations, languages and libraries. Boolean searches involving AND and OR is supported. Results can be sorted by title, publisher, publication date and other values and selected for saving to disk or e-mailing.

The search and results pages are more aesthetically pleasing than JournalTOCs, because they're less dense-text wise, it's less cramped – search results are displayed to the full width of the browser and on a separate page and partly due to the use of panels and subtle colouring to group information and buttons. Rounded edges on these add to their visual appeal. The widgets are JavaScript.

The site provides an accessibility statement which states the reasons for the visual design and how these are used (e.g. style-sheets, relative fonts etc).

7.5 Zetoc

<http://zetoc.mimas.ac.uk/>.

Mimas at The University of Manchester, JISC, British Library.

Access to the British Library's Electronic Table of Contents. It only allows searches across journal or article title, publisher, author and date, no booleans or wild-cards. Abstracts are occasionally available. Each search result record contains links to information about inter-library loans and purchasing copies from the British Library.

It said logins were required but I could run searches over it. The search results are paginated (though the total number of results are displayed). They cannot be sorted. It's more aesthetically pleasing than JournalTOCs but only because there is more white-space and it's less cramped – a benefit of avoiding the JournalTOCs three-column format. The interface is HTML forms, no JavaScript.

The site provides terms and conditions of use and quality of service information, for example:

2.2 You shall use the Records and the information available or obtained from the Records solely for the purposes of education, teaching, distant learning, private study and/or research.

5.1 Subject to Clause 5.2 and a service level agreement (SLA) between Mimas and the JISC Monitoring Unit, access to Zetoc should be available 24 hours per day 7 days per week. The SLA requires that Mimas service machines and communications facilities will be available for 99.0% of scheduled uptime and scheduled downtime of less than 10.4 hours per year. (Scheduled downtime is planned downtime for which the user community has received at least 10 working days warning. For further details see the JISC Monitoring Unit website)

7.6 How JournalTOCs could compete

7.6.1 Usability

In terms of their interfaces, my view is that the competitors are about as usable as JournalTOCs when comparing similar functionality. The ability to control how search results are presented (e.g. a list of all results and/or paginated), save search results, save previous searches and sort search results would make JournalTOCs easier to use by giving users more control over what is presented and how. Naturally, providing additional functions currently supported by its competitors would help place it in a more advantageous position e.g. supporting e-mail alerts.

The main advantage the competitors have is in terms of their aesthetic look. To varying degrees they seem less cluttered and cramped and make better use of white space. Removing the three-column layout and adopting the suggestions of 2.1 and 3.6 would go some way towards addressing this. Likewise, adoption of a JavaScript widget set (e.g. that supported by Google Web Toolkit) that offers a greater range of widgets, especially for tables, panels and buttons would further enhance its look.

7.6.2 Promoting JournalTOCs

JournalTOCs' nearest competitors are ticTOCs, which has now terminated as a project but still is maintained, and Zetoc, which, as it's active, is the closest competitor. Web of Science and Scopus are undoubtedly very powerful tools but seem targeted towards those who want fully cross-referenced publication and citation services for in-depth literature searching. Suncat is at the other end, offers only lists of journals available. So, promoting JournalTOCs could involve highlighting how:

- It is a lightweight and easy way for users to find journals of interest and subscribe to their RSS feeds to receive updates from them – it is a tool to enable users to place themselves in a position to be kept up-to-date with developments in their field.
- It is free.
- It is an ongoing project.
- It is automatically updated direct from publishers' feeds.
- It is not intended to be a heavyweight comprehensive citation and publication search service.

More generally, other ways of promoting and increasing the visibility of JournalTOCs could include:

- Asking if publishers whose feeds you expose would be willing to endorse JournalTOCs and the service it offers and provide a quote for your site and allow the use of their logo.
- Links to services e.g. TechXtra or WattJournals that have been built using JournalTOCs.
- Asking these services to link back to the web site e.g. “Powered by JournalTOCs”.
- Using consistent product naming.
- Designing a logo.
- Getting a domain name e.g. journaltocs.ac.uk. These can be purchased from as little as £3 a year.