

Sherif Scopus Enhancement Group Committee Meeting

19th February 2019
Elsevier Ltd, London

Attendees:

Sherif members:

Patti Biggs, The Francis Crick Institute - Committee secretary
Peter Bradley, University of Bath
William Farrell, University of Leicester
Elizabeth McHugh, University of the Highlands and Islands - Committee chair
Lynne Meehan, Imperial College London/University of Cambridge

Supplier representatives:

Susanne Steinginga - Senior Product Manager Scopus Content
Kai Wan - Senior Product Manager Scopus
Gillian Griffiths - Senior Product Manager Scopus
Charles Martinez - Customer Consultant at Elsevier

Apologies:

Katherine Stephan, Liverpool John Moores University

1. Welcome and Introductions

Everyone briefly introduced themselves. Lucy Ayre (University of Leicester) will be replaced Laurian Williamson on committee, but was unfortunately unwell, so William attend as her substitute.

2. Minutes of the last meeting - 26th June 2018

No queries on the previous minutes and no actions to follow up.

3. Scopus Update and Roadmap summary (see also pdf of slides presented at meeting)

Susanne reviewed what had occurred in the second half of 2018. Elsevier did not secure the contract to supply data for REF2021, but will continue to support the UK research community through its products. The REF module in Pure will continue to directly integrate with the REF submission system. The UK government will continue to use Scopus data for the BEIS UK comparative performance report (example <https://www.gov.uk/government/publications/performance-of-the-uk-research-base-international-comparison-2016>) and work with UKRI and Universities UK on monitoring the transition to open access. Additional Elsevier uses SCOPUS data for its free reports eg 'Artificial intelligence: how knowledge is created, transferred, and used' https://www.elsevier.com/_data/assets/pdf_file/0010/823654/ACAD-RL-AS-RE-ai-report-WEB.pdf and 'A Metrics-Based Assessment of Scotland's Science Landscape (2007-2016)' <https://www.elsevier.com/research-intelligence/resource-library/a-metrics-based-assessment-of-scotlands-science-landscape>.

Susanne also provided some updated statistics on new content in Scopus. Three million new items added annually. In 2018 the Content Selection and Advisory Board evaluated 2356 titles of which 1226 (54%) were accepted for indexing, they also re-evaluated 263 titles and discontinued indexing of 134 titles (51%). For more information on the criteria

used for selection and re-evaluation see

<https://www.elsevier.com/solutions/scopus/how-scopus-works/content/content-policy-and-selection>. Susanne also shared some collated data from 2016-2018 on re-evaluations. Journals may be re-evaluated if their metrics fail to meet target 2 years in a row, or if warnings are received about publication standards, or if Elsevier's Radar tool (a data analytics algorithm) detects outlier journal behaviour. In the 3 year period 772 titles were re-evaluated and 447 (58%) discontinued. On the slide pale blue circles are those selected for re-evaluation, red circles are discontinued numbers, green circles are those retained.

In the next section of her presentation Susanne looked at the ongoing work on data quality which ensures that you can rely Scopus' data for bibliometric exercises. She presented data on the volume of data items (3.7 terabytes pa) and records being assessed (8.5k per day) and information on the KPIs which are used to measure Scopus' data quality. As part of this program most records which were missing a DOI now have them (only 0.2 don't have DOI now). Work is still ongoing to improve the scores for Affiliation Organization and Affiliation Country. Data was also presented on the completeness of coverage within 18 months after start of publication year, on 1st Jan 2019 90.33% of records for 2018 published papers were complete, an improvement on same point in 2018 for 2017 records when it was 87.1%. Scopus has also been working since Aug 2017 to reduce the number of duplicate records, by Nov 2018 they were at 99.8% uniqueness. If anyone finds any duplicate records they should be reported to Scopus by giving the eIDs, the records will be merged and citation counts amended so that no citations are lost

To conclude this section Susanne mentioned some research at CWTS (Leiden University) which used Scopus as a baseline to compare WoS, Scopus, Dimensions and Crossref as bibliographic data sources. This research will be published (some of the data is available at <https://www.slideshare.net/LudoWaltman/web-of-science-scopus-dimensions-and-beyond-the-evolving-landscape-of-bibliometric-data-sources>). Scopus came out well on all their measures, despite the research using Scopus data from April 2018. Scopus has done more work in last 8 months to correct missing data.

Action: Susanne to share details of this research paper when it is published.

The next section of the presentation by Susanne looked at the 2019 roadmap. This came under 5 subheadings:

- Content selection ongoing work on title evaluation and re-evaluations of journal. In second half of year talking at other local indices, eg KCI from Korea, CSCD from China, SciELO from South America, about potentially including their content in Scopus
- Content expansion programs - Books expansion and Preprints (more details below)
- Content enrichment - Non-roman script references will be captured and translated, so can add these to citation counts for papers. This is planned for Q1-Q3
- Metrics - this is the subject of Kai's presentation later in minutes
- Content quality - Data quality programme will be ongoing. In the middle of the year will look at Flexible Source Browse and Flexible Multiple Subject Classification for journals

Preprints - Scopus are investigating adding preprints during 2019 as part of the growth strategy as they increase currency of research and are growing quickly (30% compared to article growth rate of 2-3%) annually over past two years. There are 4 major preprint servers which Scopus would like to index fully - ArXiv, SSRN, BioRxiv and ChemRxiv. As preprints are not peer reviewed questions have been asked about how to ensure the quality of preprints. The second slide on Preprints present the 6 key questions that Scopus

will be considering at a meeting shortly. The meeting felt that preprints should be kept separately from full articles. Current thinking is that citation counts for preprints and papers should be kept separate too. Content licencing negotiations and technical discussions are both ongoing

Under Data quality a key area will be improving the precision of the funding data presented in Scopus. The information is extracted from the funding text by Machine Learning tool and presented as a simple table. A new release of Machine Learning tool should be better at capturing this information and Scopus will be re-indexing papers back to 2008. This process will take c 6 weeks and new data is expected to be visible in Scopus in April.

The book expansion is aimed at addressing a perceived content gap in Social Sciences and Arts & Humanities, based on demand from customers, although the CWTS study found more books in Scopus than any of the other bibliographic databases. Scopus will backfill current coverage, add more publishers and include non-English language content.

Kai then present about metrics, particularly Topic of Prominence. This metric is already familiar to SciVal customers, but is now going to be surfaced in Scopus in both articles and author profiles. It is an indicator of the current momentum of a topic. The article level indicator is a calculated weighting of 3 metrics for papers compared other paper in the topic area:

- Citation count in a year compared to other papers in same and previous year
- Scopus view count in a year compared to other papers in same and previous year
- Average Journal CiteScore for the year

For articles the SciVal Topic Prominence will appear as block below the abstract.

Kai also reported that there will be changes to the way funding details are displayed. At present each grant is presented a single line, in future there will be a single line per funder with multiple grants. This change will also link to central Elsevier funding data which is manually curated for accuracy.

Gillian presented about the Institutional Profile Wizard. This allows institutions to see and adjust their profile data held on them. More information on the wizard is available at <https://blog.scopus.com/posts/scopus-institution-profile-wizard>. You need to register to use this wizard as it gives administrators the power to amend structure of an institution which will be displayed publically. The wizard only shows the current hierarchy, not historical data. The alternative names are clusters generated by algorithm. Administrators can look for unmatched clusters of names and associate them with their institution.

4. Queries and Comments from Sherif members

a. Follow-ups to matters raised in previous meeting.

1. Kings College London comments on author profiling and cruciality of this to next version of Pure was noted. The comment on ensuring that the top menu remains visible was also noted and will be feedback to the design team. Kings are working around this with SciVal integration

Action: Kai to report on design team response

2. The Open University reported a book record disappearing from Scopus, this content has been restored. More generally they were concerned about the

helpdesk's misunderstanding of the issue they had raised and the resulting inappropriate responses, even when they went back to the helpdesk. Others around the table had also experienced similar issues with the support desk. Susanne responded that these issues picked up by monitoring and additional training provided to the helpdesk staff.

Response: This was a misunderstanding by helpdesk team. Any problems with 1st line support should be referred up to 2nd line support and if this fails to resolve the issue it can be raise with your customer consultant:

- Charles Martinez - South of England
- Matthew Walker - North of England
- Denis Reidy - Wales, Scotland and Northern Ireland

Action: Kai to ask if the support KPIs can be shared with this committee.

3. The Open University also raised an issue with exporting records from Scopus to Mendeley. A generic error message appears which doesn't indicate that the user needs to sign in to their Elsevier account and associate this with their Mendeley account. This is a problem on the Mendeley site due to the new logon system and non-merger of accounts.

Response: this has been fixed.

4. The University of Bath had raised five points. Kai responded to their request for CSV exports to include Scopus Author ID, the API would be a better way to get this data. There will be an expansion to the number of records that can be viewed and downloaded in third quarter of 2018. The request for information on the sources of patents in Scopus was noted. Translations of patents are not available, but users can discover what the patent is about by looking at others in the same patent family. The broken links to Espacenet patents have been fixed. With regard to the request for more information on sources behind data search and future plans for it, Kai has forward this to Mendeley for answering.

Response: Scopus Author ID has been added to CSV exports. Kai reported that number of downloads will be increasing by end of Q2 2019. Currently 2000 full records can be download, will increase to 5000. Currently 20,000 citations can be downloaded, this will increase to 50,000. On patents Kai reported that a field has been added for Patent Office so can filter on Patent Office. More can be added without needing to do full reindex. Numbers of patent offices included are being expanded from 8 to 13, this is in progress with no known end date. The query on Mendeley Date needs to be referred to specialist team.

Actions: Kai to refer Mendeley data query to the specialist team.
Kai to ask product manager for a list of the patent sources.

5. The University of Cardiff raised two matters. The first was a search containing an OR between two phrases. The Boolean OR is being action first then the implied ANDs. Not possible to change the logic, but might get better results if used quotations or brackets around the phrases. The comment on combining searches being less intuitive than other discovery tools was noted.

Response: Acknowledgement that search needs to be more intuitive. The Advance search - outline query matches brackets. Advice was to use brackets with AND NOT at end of search string. Also good to use help files as search syntax is different from WoS.

6. The University of East Anglia made a similar comment about combining searches. They also find users confused by the two step process to use personalisation and requested that Scopus look at simplifying it, also making the Login button more prominent than Register. This feedback will be shared with the interface design team. They had also commented on the intrusive number of webinar popups which occurring training sessions last year: Kai has provided this feedback to the appropriate team.

Response: Changes to the interface - a new product manager advised to look at making it more intuitive without losing functionality. The loss of ability to add extra terms has also been noted. Webinar popups - these must stay, but if dismissed no new popup should appear.

Additional comment from Lynne - Personalisation breaks part of functionality.

Action: Kai to have a further conversation

7. The University of Plymouth had noticed some inconsistencies in the indexing of book chapters, Susanne undertook to raise this and provide an explanation.

Action: Susanne thought she had sent notes, but will resend them.

8. As we looked at the live demos a number of additional comments were made. On the user interface some of the links are very small eg Sources and Related Sources - Susanne note this feedback. A user of Mendeley comment that it produced an overwhelming number of recommendations and social interactions which made it difficult to use for citation management - Kai will pass this feedback to them.

Response: Small links - planning to redesign pages. Mendeley - the feedback was passed on, but the individual who was handling this has left the company.

9. The Francis Crick Institute added a couple of additional matters. Clicking on titles in search results giving a null results page the first time they are click. This seems to be linked to a problem with individual accounts which are trying to connect to phantom Refworks accounts. The user removes the erroneous Refworks details, but they keep repopulating. Charles M commented that he also experiences this problem and that it is under investigation. The Crick had also notice that in its alert searches for new publications we receive some results for Cell Press papers which are "Article in Press" are not yet in visible in Scopus or ScienceDirect. Generally they appear in Scopus about 3-4 days after our alert search. Susanne will investigate if PB can send her some examples.

Response: Refworks issue - it should have been fixed in Dec, but is still occurring. Alert issue with Articles in press - has been resolved.

Action: Kai and PB to take up Refworks issue outside the meeting.

b. New feedback from sheriff institutions to this meeting.

Peter B asked a couple of question on behalf of University of Bath:

10. Where is SCOPUS with full text reader? Kai responded that it was in progress, but he could not share the time line at present.
11. Can the citation count include citations for article when it was in press. Kai thought that they did carry forward to full article.

Action: Kai to double check on citation counts for article in press.

Appendix A gives some feedback which Scopus had requested at the last meeting, it also contains some questions which had been submitted prior to the meeting. Elizabeth apologised that she had forgotten to send a requests for questions and queries to lis-e-resources this time, so we only have relative small number of responses. She will also encourage the inclusion of screenshots of problems as this is easier to understand than long texts.

Feedback from University of Bath:

12. Their response to Scopus' question clearly rolls in our earlier discussion on preprints. Hopefully this will be clearer for the next meeting
13. Timeliness of indexing is difficult to judge as don't have comparative data for other companies, only SCOPUS makes its data available.
14. The responder seems to have misunderstood the question about whether it would be useful for Scopus to links to green copies (ie Accepted manuscripts) in repositories as these are different to preprints.
15. The responder didn't seem to support Full text reader. Peter added that if it isn't available people can use open URL resolvers to access full text, or use the download facility in Scopus. However the download facility does always pick up content available though EBSCO host, so users may not realise that they have access. Is may be down to a broken triangle between discovery tool, aggregator and journal. Kai responded that a work around would be to send a list to second line support of Journals and ISSNs, so that the link resolver button only appears on where you have access. Hopefully this problem will improve with new download tool.

Feedback from University of Cambridge/Imperial via Lynne

16. Updates issue was due A-B testing. Suggested that a manual button be provide to change back to the stable interface as it can be difficult if during a training session all the class aren't seeing same thing. Re the request for more information on updates there will be information on latest updates on the landing page, but irregular used may miss these, so the Scopus blog can be useful as it describes all changes. Bimonthly catchups with your custom manager can also be requested
17. Combining results -. The request for a better way of combining searches was noted.
18. Advanced search - the query about search for publishers could be overcome by searching on DOI handle followed by wildcard eg 10.1038* or adding Publisher(*name*) to search strategy

Feedback from Edinburgh University:

19. Had asked that Limits be included in Scopus' search history. Scopus had responded to this that in the facets panel you do have ability to search with the results and that the refined search will appear as a new entry in your search history, so this can be used to add words and limits to a search.

Feedback from Leicester University:

20. The comment about keywords did not have sufficient detail to understand the issue or concern. William did not know the context of this comment either.
21. The question about Institutional dashboard was answered. EPIC dashboard includes all the request sources except SciVal.
22. Request for information on merging authors was answered. The user or administrator used the Author wizard, completes 3 steps then submits this. This creates a ticket which the user can track via the dashboard (visible as hamburger icon when you've signed in). Requests to merge records are always manually review by Scopus staff. Charles reminded everyone that the dashboard tracks all tickets for Scopus and SciVal in a single place.
23. Question on search multiple affiliations for an author was answered. Searching multiple affiliations in Author search is not possible as the affiliations are held in separate indices from names. Kai commented that a colleague (Leonora) is investigating allowing author profiles to have multiple affiliations - this may affect search too.
24. The query about able of identify Embase material in Scopus was answered. Charles responded that in advanced search there is a field called INDEX, so you can include Embase using AND INDEX(embase) or exclude Embase using AND NOT INDEX(embase).
25. Desire for unlimited search terms and line. Sorry this is not possible as there are technical limitation on resource handling.

Having completed the pre-notified questions we asked around table if there were any further issues to raise. Elizabeth from University of Highlands and Island had two items:

26. When will new version of Refworks be integrated into Scopus? Unfortunately there has been no response from Refworks to Elsevier's request to use their API.

Action: Elizabeth to refer this the Refworks Enhancement group as it needs co-operation from the supplier.

27. Her university is have some difficulties seeing Scopus articles in their discovery service Encore Duet unless they are logged in to the discovery service. For other products they do not need to login. The response is that they need to obtain a Scopus API key for Ebsco for this to work. They will also need to email Integrations support team (integrationsupport@elsevier.com) to request an exemption from the fair use policy for data transfer.

5. Group Discussions

The last 4 slides of the presentation introduced discussions on Preprints, additional Indices, institutional profiles and analytical tools in Scopus

Preprints: Scopus captures citations of preprints in papers (inbound), but it would be valuable capture citations of journal in preprints (outbound)? This seemed valuable to journals, but we felt that it might be best to keep these in a separate count. Preprints might also feed into topics of prominence as they are at the cutting edge of research. A second question was should preprints appear in author profiles. We were in favour of this provided they were listed separately from full articles and the author choose whether or not to show them. This would also make preprints available in the API feeds to CRISs.

Local indices: The question was whether adding these as separate databases was a good idea? The feeling was that more international, quality assessed sources were a good thing. The University of Leicester has a lot of Chinese students and so the inclusion of CSCD might encourage them to use Scopus. Peter raise a concern about the ability to deliver documents discovered from these additional source - universities would need to manage their users expectations.

Institutional Profiles: Gillian who is leading on this area had a series of questions, currently the profiles only reflect now. Many institutions change as they incorporate small institutions or spin out parts of institution as separate entities - if Scopus was to reflect these changes institutional hierarchies might get very tangled. Would adding dates help with this. The committee though this was a very complex issue as it is different organisation to organisation. Linking dates to entities would probably help. The question about attribution of publications after a change is particularly difficult as it might depend on the legal agreement reached between the organisations involved.

Analytical tools: Kai had three points for discussion. His first was alternative graph views for the analytical tools as these have not been reviewed for a while. The committee would like to see better tools in Scopus rather than having to export data to programmes like VoS Viewer. Linear information with cross links is an easier sell to data users. His second point was that the Sources page in Scopus is going to be redesigned, proposal is to add "save journal list" function to the page and make this exportable/shareable. This could be linked to an alert of this saved list. The committee had no objections to this proposal. His final point was simply a request for any general feedback on Source page search and results functions. None was forthcoming on the day, but we will share this request more widely.

Action: Share request for any general feedback on Source page search and results functions

6. AOB

Charles commented that the current Scopus licencing agreement comes to an end in March 2020, so negotiations for an new licence agreement are starting, He hopes the merger of Eduserv and JISC will lead to less oconfusion and even the possibility of compound deals for multiple Elsevier products

7. Date and place of next meeting

Date and venue to be confirmed as needs to fit between Elsevier quarterly meetings.

Appendix A: Scopus Feedback from sheriff Institutions: February 2019

University of Bath

Feedback from: Elsevier's request (item 5; p.5 of the minutes):

Q1. What would be most valuable & useful to you?

Q1: I think I'd need a bit more clarification what they mean [by] the distinction between "all scholarly content" and "curated, high quality, peer reviewed content" to judge whether it'd be valuable to expand current coverage. I'd have thought that a key benefit of Scopus as opposed to e.g. Google Scholar is its collection selection policy and the quality of content indexed. I'd think, with a clearer definition of this established, we'd probably want to have the ability to include/exclude.... I wouldn't like to see this adding to subscription costs. There is already the facility to find some patent results in Scopus..."

Expansion in coverage, given Elsevier's stringent inclusion policy: with or without a filter-and filtering the curated and peer-reviewed from the non-curated: please clarify this further, explaining in detail.

Q2. And what about: currency/timeliness? How important is timely indexing to you? What indexing time (time from publication to present in product) is acceptable to you?

Q2: "The more timely the indexing the better. Indexing time for different databases is not really something I've a clear picture of, so it might be useful to get more information on. I think this has come up occasionally in enquiries e.g. from PhD students in DocSkills.

From another colleague: "I would like to see more indexing of 'online first'/'early online'/'in press' articles. It's not unusual for articles to be published as 'online first' for a year [+] before they are assigned a publication volume etc, so waiting for the final publication detracts from currency.... My understanding is Scopus index articles in press if the publisher gives them the details, but they don't routinely aim to index all 'in press' articles for the journals they cover."

Q3: Green OA indicators:

- a. What is Scopus would extend its OA coverage with 'Free at Repository' content.
- b. This will allow users to still have access to the content even though it may not be the best version of the content (which would be 'Free at Publisher')
- c. Would this function benefit your use of Scopus?

Q3: "I think this would be useful, so long as it is clear that these are pre-prints"

Q4. Scopus full text reader:

- a. This function allows users to download and display full text articles in Scopus.
- b. Would you prefer to read the full text in Scopus or is it OK to download it first from publisher website before you can read it?

Q4: I can't see any particular benefit to this for us, but I might be missing something!

University of Cambridge

Suggestions

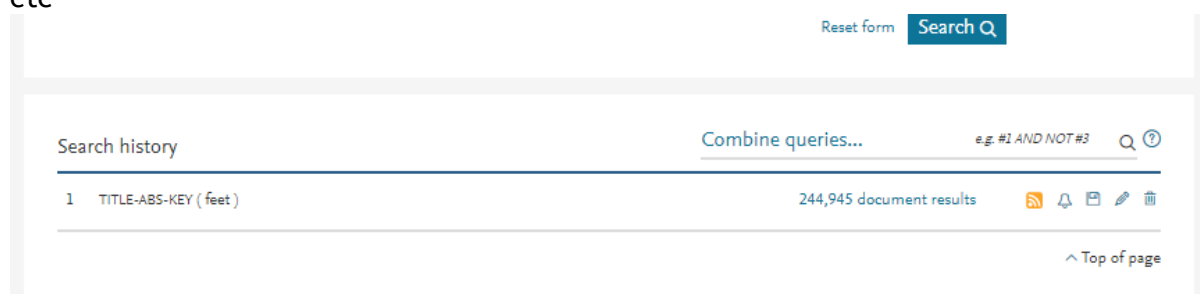
Updates

If you could ask them to stop slightly altering things that would be great! There was a month where you had to login in order to be able to export things but luckily that seems to have stopped.

Also, informing libraries of updates – if we don't use Scopus regularly an update can take place and we won't know what has changed. Can we sign up to a mailing list to keep us up to date.

Combining Results

If you have to combine lines using #1 etc, please can the line numbers be identified as #1 etc



Lines are currently just 1, 2 and so on.

Another alternative was to use tick boxes

My big thing is to ask if they could have tick boxes or something similar in the search history. It's such a palaver having to type #1 or #2 or #3 (etc) when you want to combine your searches.

Advanced Search

Used to build queries to search Scopus for content based on funding, publisher, OA status, and data access statements.

One annoyance is that when building a search query in the advanced search I can restrict my search to a specific journal, (SRCTITLE) or (EXACTSRCTITLE), but I can't restrict at the publisher level, which we'd want to be able to do.

Also, when searching for OA content, the current field code, (ACCESSTYPE), is not sophisticated enough. We need to be able to distinguish between gold papers in fully OA journals, gold in hybrid, and green OA articles via repositories. WoS have similar functionality. We'd also want to be able to search by licence, to restrict our results to only items with a CC-BY licence, for instance, or CC-BY-NC, CC-BY-NC-ND, publisher own Gold licence, etc.

Finally it would be very useful if it were possible to add a field code related to data access statements, so that we could limit our search to items with/without a data access statement, and for the statement to be pulled through into the export of results.

Praise for Scopus

just to say that I use it extensively for my PhD research, and I have used features in terms of advanced searching, exporting docs, and exploring citations extensively, and it hands down the most robust database I've ever used. When I have to use a ProQuest or Ebsco database now, the experience is horrible in contrast (e.g., you can only export a page of search results at a time in ProQuest databases!).

University of Edinburgh

A request for development go to Elsevier such that Limits can be applied to Scopus's Search History searches at the time of combination. Not just via refine panels at results stage.

e.g.

Combine queries...
#1 AND #2 AND #3 e.g. #1 AND NOT #3

Search history	Document results	Actions
3 TITLE-ABS-KEY (conceptual* OR perspective* OR "world view" OR narrative*)	1,482,725 document results	[Share] [Alert] [Export] [Edit] [Delete]
2 TITLE-ABS-KEY (resilien* OR "mental health" OR wellness OR "well being" OR wellbeing OR strengths OR psychosocial OR "Protective factor*" OR "coping behavior*" OR "support system*" OR "risk factor*" OR trauma* OR "post-trauma*" OR posttrauma*)	3,924,435 document results	[Share] [Alert] [Export] [Edit] [Delete]
1 TITLE-ABS-KEY (indigenous OR native OR aborigin* OR "Pacific Islander*" OR maori* OR "American Indian*" OR "First Nation*" OR inuit* OR metis)	528,544 document results	[Share] [Alert] [Export] [Edit] [Delete]

I'd like limit options here, eg 2000 onwards, review articles only, English language.... that then apply to the combination

^ Top of page

Rather than doing the search and then using the refine:

Refine results

Limit to Exclude

Access type ⓘ

- Open Access (384) >
- Other (1,933) >

Year

- 2019 **ticking all etc** (21) >
- 2018 (249) >
- 2017 (248) >
- 2016 (220) >
- 2015 (195) >
- 2014 (187) >
- 2013 (182) >

University of Leicester

- Keywords generated by learning algorithm (IEEE & Archive of Virology).
- Institutional usage dashboard for Scopus, SciVal, ScienceDirect & Clinical Key.
- Author merge requests: who does these, what is the process?
- Author search: ability to select more than one affiliation e.g. David Clark (Leicester & Oxford).
- List of Embase material in Scopus for systematic reviews. Search Embase separately/facet Embase records off.
- Maximum number of search terms in a line/lines combined together. No limit!