Original Article

A Survey of the Current Status of Web-Based Databases Indexing Iranian Journals

Shahin Merat MD•*, Shahab Khatibzadeh MD*,**, Bita Mesgarpour MD***, Reza Malekzadeh MD*

Background: The scientific output of Iran is increasing rapidly during the recent years. Unfortunately, most papers are published in journals which are not indexed by popular indexing systems and many of them are in Persian without English translation. This makes the results of Iranian scientific research unavailable to other researchers, including Iranians. The aim of this study was to evaluate the quality of current web-based databases indexing scientific articles published in Iran.

Methods: We identified web-based databases which indexed scientific journals published in Iran using popular search engines. The sites were then subjected to a series of tests to evaluate their coverage, search capabilities, stability, accuracy of information, consistency, accessibility, ease of use, and other features. Results were compared with each other to identify strengths and shortcomings of each site.

Results: Five web sites were indentified. None had a complete coverage on scientific Iranian journals. The search capabilities were less than optimal in most sites. English translations of research titles, author names, keywords, and abstracts of Persian-language articles did not follow standards. Some sites did not cover abstracts. Numerous typing errors make searches ineffective and citation indexing unreliable.

Conclusions: None of the currently available indexing sites are capable of presenting Iranian research to the international scientific community. The government should intervene by enforcing policies designed to facilitate indexing through a systematic approach. The policies should address Iranian journals, authors, and indexing sites. Iranian journals should be required to provide their indexing data, including references, electronically; authors should provide correct indexing information to journals; and indexing sites should improve their software to meet standards set by the government.

Archives of Iranian Medicine, Volume 12, Number 3, 2009: 271 – 278.

Keywords: Indexing and abstracting • information services • Iran

Introduction

During the recent years, with the rapid growth of universities and research centers in Iran, a growing number of scientific articles are being produced.[1–3] The majority of these articles are published in Iranian journals, many in Persian, which are not indexed in any international system. This literature is extremely difficult to access and is thus rarely, if ever, cited. Articles published in Persian are rarely available to non-Iranians. Even when English translations of Persian abstracts are present, they are only found on search engines such as Google™ or Yahoo® and not generally available on international scientific indexing systems. The visibility of research articles is extremely important to the scientific development of the country. In order to increase visibility, most Persian-language journals require the authors to provide an English translation of the abstract and

Authors' affiliations: *Digestive Disease Research Center, Shariati Hospital, Tehran University of Medical Sciences, **Center for Academic and Health Policy, Tehran University of Medical Sciences, ***Scientometry Department, Deputy of Research and Technology, Ministry of Health and Medical Education, Tehran, Iran.

•**Corresponding author and reprints:** Shahin Merat MD, Digestive Disease Research Center, Shariati Hospital, North Kargar St., Tehran 14117, Iran.
Telefax: +98-218-801-2992, E-mail: merat@ams.ac.ir
Accepted for publication: 4 April 2009

many others provide the full text of their articles on their web site. Although these measures help, they are by no means a replacement for a good indexing system. It is an absolute requirement for the scientific community of Iran and other Persian-speaking countries to have at least one well-designed, all-inclusive, indexing system covering scientific articles published in these countries.[4]

A good indexing system should have many characteristics. English interface, advanced search on various fields, and user-friendly interface are basic. The coverage should be complete and include all Journals approved by the Ministry of Science, Research, and Technology (MOSRT) and the Ministry of Health and Medical Education (MOHME). Citations should also be indexed and be searchable with the possibility of calculating impact factors and other scientometry indices.

The first step would be studying the current status. Some private and governmental organizations have attempted to index articles published in Iranian journals. Many provide English abstracts and various levels of search capabilities. But there is no systematic evaluation of the currently available indexing systems. In this study, we attempt to systematically evaluate the various web sites claiming to index scientific articles published in Iran.

Materials and Methods

Identifying current indexing sites

To identify existing indexing sites, the well-known indexing sites were listed by asking researchers and experts. Additionally, two frequently used search engines, Yahoo® and Google™, were used. The search was performed using the Persian words for "journal" or "publication" or "magazine" and "database" or "bank" or "information" or "guide" or "citation report" and "Iran" or "Iranian".

Sites claiming to index journals or articles published in Iran were included. Exclusion criteria included indexing less than 50 journal titles, indexing only specific subjects, and not being updated during the last 12 months.

Identifying scientific journals published in Iran

All scientific journals are required by law to be registered by either MOSRT or MOHME. Thus the list of approved journals was requested from these two ministries.

Determining the coverage of indexing sites

We contacted the administration of each web site and asked for the number of journals and the number of issues indexed for each journal. Where possible, the data were also extracted from the web site itself and checked against the report from the administration and inconsistencies were resolved. The journals indexed were categorized into those registered by MOHME and those registered by MOSRT. To find the coverage of each site for each journal, the number of published issues should be known. Unfortunately due to the large number of journals involved, it was not possible to find the exact number of published issues. So, in order to express the coverage, the number of issues of each journal indexed by each site was compared against the largest number of issues of that journal indexed by any site. Thus, if the coverage of a site for a particular journal is 100%, it means that among the sites evaluated, that site has indexed the largest number of issues of that journal.

Evaluation of content and accuracy

The journals were sorted on subject and then alphabetically. A random starting point between one and ten was chosen and then every 10th journal was selected as sample. If the selected journal was not indexed in all sites, the next 10th journal was selected. From each journal, two issues that were indexed in all sites were selected, and from each issue one article was randomly selected. The completeness and accuracy of the data entered in each site for these articles were studied.

Accessibility

The accessibility of the indexing sites was compared. The factors studied as indicators of accessibility included: access limitation, registration, stability of site address, internet connection speed, online help, and availability of English access.[5,6]

Search and advanced search options

Options studied included the possibility of using logical operators (and, or, not, …), combining search in different fields, limiting search results, advanced search, etc.[5,6]

The sensitivity and specificity of the search engine were evaluated by 10 single-word searches and 10 combined searches. The number of results found is a rough indicator of sensitivity. This number is affected by the number of issues indexed by that particular site. So, we also reported the

number of search results per issue as a better indicator of sensitivity.

The percentage of relevant results was reported as an indicator of specificity of the search engine.

Currency of data

The publishing date of the latest issues of 45 journals (selected earlier) was inquired from the journals' editorial office. The web sites were then followed for the next three months to see when the issues appeared.

Rate of addition of new articles and usage

The rate at which new articles were added and the usage was inquired from the administration of each web site.

Citation index

The presence of citation index and the possibility of extracting citations were evaluated in each site.

Results

Identifying current indexing sites

After applying the inclusion and exclusion criteria to the list of sites found by the search engines, five web sites were found which claimed to index Iranian research articles (Table 1). IranMedex only indexes journals registered by MOHME and IranDoc only indexes MOSRT journals. The search was performed in September 2007.

Identifying scientific journals published in Iran

MOHME has registered 122 journals and MOSRT 269 as of October 2007. A total number of 391 journals registered and published in Iran were studied.

Determining the coverage of indexing sites

The coverage of each journal for each site was calculated as explained in methods. The mean coverage of each site calculated separately for journals registered by MOHME and MOSRT in January 2008 are given in Table 2.

Evaluation of content and accuracy

As described in methods, 45 journals and 90 articles were selected. Four articles were not used since they were not listed in most sites.

The number of issues in content and accuracy of the 86 sample papers in the five sites are too many to be detailed here. Results are presented in general with occasional references to individual sites.

Corresponding authors and their addresses were available in only 60% of papers in Scientific Information Database (SID) and none of the other sites. English abstracts for the Persian language papers were totally unavailable in MagIran and missing in about 30% of cases in other sites. Shiraz Regional Library of Sciences and Technology (SRLST) provides no abstracts (whether English or Persian) although a nonsearchable full-text is available for some papers which is accessible through another database.

In most sites there are separate pages to search English and Persian papers which are usually not easy to switch.

The English translation of titles and authors of Persian language papers were not available in all cases and in all sites. In SID, these data, including the English abstract, was only found in the English section of the site which was totally separate from the Persian section. In IranMedix, English and Persian information were available on the same page. The English translation of dates was not converted to Georgian calendar in IranMedex. In SID, one abstract was totally unrelated to the title and had probably been uploaded incorrectly. In SRLST when the number of authors exceeded three or four, only the first author was listed followed by "et al." and searches on all authors except the first failed. Furthermore, in some papers authors were listed as translators.

Typing errors were seen in all sites. The largest number of typing errors was seen in SID with at least one error in 59% of selected articles, followed closely by SRLST with errors in 52% of articles. Errors included typing errors in title, authors,

Table 1. Web-based databases indexing Iranian scientific literature

IranDoc	http://www.irandoc.ac.ir/
IranMedex	http://www.IranMedex.com/
MagIran	http://www.magiran.com/
Scientific Information Database (SID)	http://www.sid.ir/
Shiraz Regional Library of Sciences and Technology (SRLST)	http://www.srlst.com/
(now Regional Information Center for Science and Technology)	(now http://www.ricest.ac.ir)

Table 2. Coverage of Iranian indexing sites

	Mean coverage of MOHME journals	Mean coverage of MOSRT journals	Total number of issues indexed
IranDoc	NA	21.3%	2495
IranMedex	82.0%	NA	2927
MagIran	38.7%	27.8%	3550
SID	70.6%	36.1%	5331
SRLST*	72.8%	90.1%	9615

NA= not applicable; *Data were collected seven months later than other sites in August 2008. We were not able to confirm these data.

keywords, abstracts, incorrect order of authors, missing authors' names, inconsistencies in authors' name format, inconsistencies in capitalizations, punctuation errors, errors in page numbers and dates, missing end-page numbers, and missing words in titles.

Another important problem was inconsistent and often arbitrary translation of Persian journal titles to English, even within the same sites, making citation indexing impossible.

Accessibility

The accessibility items checked for each site are given in Table 3. The check was performed in January 2008.

IranMedex and MagIran are run by non-governmental organizations and require paid subscription or contracts to allow access. SRLST also requires subscription. SID and IranDoc have no access limitations although IranDoc requires a registration which is free and fast. SRLST has the language options of Arabic and Kurdish, but these options only apply to the interface and do not provide Arabic or Kurdish translations. SRLST has advertised its address as www.srlst.com, www.rlst.ac.ir, and http://www.ricest.ac.ir at different times; but it is currently accessible only at http://www.ricest.ac.ir. During October 2007 through March 2008 the web site of SRLST was highly unstable and we were not able to study its features. We were not able to get convincing answers from the help desk and administration until May 2008 when we were told that the instability was due to hardware failures and upgrades which were then over. The online help on SRLST was inactive during the time of study in August 2008.

Search and advanced search options

A comparison of available search options in January 2008 is given in Table 4.

There is an advanced search option in SID but was not functional at the time of study. Search in abstract was not possible in SRLST. None of the sites had options for sorting search results or proposing spelling corrections or alternate search terms.

The specificity of single word search was assessed by searching for 10 words (six Persian, four English). Ten combined searches were also performed using “and” and “or” logical operators (seven Persian, three English, one mixed English and Persian). The advanced search option in SID was disabled and mixed English and Persian search was not possible in SRLST. The results returned by each search and by each site were examined to find relevant results. The results of the searches and the mean percent of relevant results are given in Table 5.

Currency of data

Unfortunately, none of the journals provided us with reliable information and thus we were not able to evaluate the currency of data entry in the five sites.

Table 3. Accessibility of Iranian indexing sites

	IranDoc	IranMedex	MagIran	SID	SRLST*
Access limited by contracts		X	X		X
Registration required	X				
Easy registration (if required)	X				
Easy loading of home page	X	X	X	X	
Search, online help, and clear instructions on home page	X	X	X	X	
Accessible in English		X		X	X
Stability of address	X	X	X	X	
Helpful and responsive help desk	X	X	X	X	
User friendly interface	X	X	X	X	

*Data were collected seven months later than other sites in August 2008.

Table 4. Search options in Iranian indexing web sites

	IranDoc	IranMedex	MagIran	SID	SRLST*
Availability of logical operators	X	X	X		X
Subject search	X				X
Keyword search		X		X	X
Title search	X	X		X	X
Phrase search	X	X	X		X
Author search	X	X		X	X
Possibility of limiting search		X			X
Advanced search		X	X		X
Possibility of formulating search	X	X			X
Reporting number of results	X	X		X	X
Option of number of results displayed per page		X			X
Option of selecting search limitations	X	X			X
Predicting search words during typing	X				
Possibility of saving search terms for later use		X			
Highlighted search terms in results				X	
Reporting search time		X			
Possibility to select fields to be printed				X	
Reporting size of retrieved documents		X			

*Data were collected seven months later than other sites on August 2008.

Rate of addition of new articles and usage

The rate of addition of new data to each site is given in Table 6. MagIran indexes nonscientific magazines and even newspapers. According to their administration, it was not possible to separate the data for scientific material; so, we do not have usage data or rate of data entry for this site. IranDoc could not provide search and view data. The usage of the sites, including number of searches, views, and downloads, is given in Table 7. It should be noted that the data presented are as reported by site administrations. We could not assess the validity of these data.

Citation index

Citation report is only provided in SID and SRLST. According to the administration of SID, all citation index data are available. But many links on the site were inactive at the time of study and only the journal citation reports were accessible which included the total number of citations, immediacy index, impact factor, highly cited articles, highly cited researchers, and highly self-cited researchers. The citation report in SRLST also provides journal impact factors.

Discussion

During this study, we identified five web-based indexing systems which index scientific articles published in Iran (Table 1). We compared various features of these sites in order to identify their strengths and shortcomings.

SRLST has the largest number of issues indexed and the best overall coverage of scientific journals. IranMedex has the best coverage for MOHME journals. It should be noted that these results are relative to other sites; otherwise, none had 100% coverage of MOHME and MOSRT journals.

Accessibility is another important feature of an indexing system. It is understandable that a non-governmental organization would depend on paid submissions in order to survive. But such a feature would make the system highly unacceptable for introducing Iranian literature to the international community; a main objective of such indexing systems.

Another main objective of such sites is search capability. The best search engine by far was IranMedex which had most of the desired options

Table 5. Results of simple and combined search in Iranian indexing web sites

	IranDoc	IranMedex	MagIran	SID	SRLST*
Simple search					
Total number of search results	560	1145	971	717	1086
Number of results per 100 issues indexed	22	39	27	13	11
Percentage of relevant results	32%	88%	55%	89%	51%
Combined search					
Total number of search results	81	8841	105	—	59
Number of results per 100 issues indexed	3	302	3	—	1
Percentage of relevant results	77%	89%	66%	—	72%

Table 6. Rate of addition of new papers to Iranian indexing sites

	2001	2002	2003	2004	2005	2006	2007 (first half)
Number of abstracts							
IranDoc	3960	8186	9352	9825	8639	7268	3200
IranMedex	—	—	11739	17051	23062	28508	31610
SID	37690	—	—	15436	63930	105983	139623
SRLST		39501	38000	25260	21498	53342	25900
Number of full-text papers							
IranDoc	—	—	–-	—	—	—	—
IranMedex	—	—	1097	6291	9885	13230	14857
SID	—	—	–-	1095	11918	25705	33624
SRLST	20602	38406	55807	86502	107100	128700	140188

of an advanced search system (Table 4). Furthermore, the specificity and sensitivity of this search system was best among the sites evaluated (Table 5). It is seen that in spite of its very good coverage, SRLST has one of the least sensitive search systems. This is expected since SRLST does not include abstracts in the search because they link full text PDF which is not searchable.

The rate of addition of new data is highest in SRLST and SID and the highest usage was in SID followed by IranMedex. There are some inconsistencies though. For example, the number of downloads are greater than the number of views and searches in SRLST. Unfortunately we could not confirm the usage data or rate of data entry.

The completeness of entries is another important issue. It appears that there is little attention to indexing corresponding authors and affiliations. English abstracts were missing in many instances and English translation of titles, keywords, or author names of Persian papers were also not generally available. This would make the Persian paper not visible to an English search; the most important objective of an indexing site. Inconsistencies in data entry and punctuations were also ubiquitous. The order of first and last name of authors varied, the capitalization of titles was non-consistent, use of punctuations was almost always incorrect, etc.

Another important problem was inconsistent and often arbitrary translation of Persian journal titles to English, even within the same sites, making citation indexing impossible.

The number of typing errors was also unacceptable. Such errors not only make searches ineffective but seriously damage the image of an indexing site. With so many errors, citation index is totally unreliable.

Although it is obvious that a great deal of planning and management effort has gone into designing and maintaining these sites, none hold up to the standards required for an international-class indexing system. Not all the responsibility of these errors can be placed on the indexing systems. Keeping track of the publication of almost 400 journals is extremely difficult and resource consuming. Entering the information and translating appropriate fields is much more demanding. In many instances, errors made in the original paper should also be handled, a very technical issue requiring well-educated editors. Entering references in order to create a citation index is almost impossible without numerous typing errors. In many cases, the correct format of

Table 7. Usage of Iranian indexing sites

	2001	2002	2003	2004	2005	2006	2007 (first half)
Number of searches							
IranDoc	—	—	—	—	—	—	—
IranMedex	—	—	—	64988	131086	207728	128539
SID	—	—	—	52114	104225	216441	449945
SRLST	81345	83576	86539	80757	90359	94893	48901
Number of views							
IranDoc	—	—	—	—	—	—	—
IranMedex	—	—	—	51397	166187	264996	182977
SID	242035	—	—	1050000	2170401	6995547	7299270
SRLST		292516	311541	242271	307220	332125	176043
Number of downloads							
IranDoc	9600	8483	9254	7263	4241	2400	1100
IranMedex	—	—	—	9705	39218	77601	59340
SID	—	—	—	13000	34805	160373	268499
SRLST	118791	232197	272690	305145	394918	444175	228703

Table 8. Information required for each journal issue published

	Comments
Journal title in English, year, volume, issue, publication date	Dates in Gregorian calendar
Epub date	Optional
Paper title in English	Format should be standardized, provided by author
Authors in English (first name, last name, affiliation)	Provided by author
Authors in Persian (even for English-language papers)	Provided by author
Corresponding author (address, e-mail, telephone, fax)	In English
Type of article	Original, review, letter, …
Short title	Optional
Keywords in English	From MeSH
Abstract in English	Limited number of words
Language	English/Persian/…
References in English	Standard format

references is not observed even in the original published article. Sharing the responsibility and creating an efficient mechanism is a must if we want to create a lasting international-class indexing of Iranian scientific journals.

A large number of research articles in Iran are published in Persian and are neither nationally nor internationally visible and thus, would not be cited. Even English articles are rarely optimally indexed on the web. This is an important shortcoming for the scientific community of Iran and a major barrier for contribution of Iranian scientists to the international scientific literature. In spite of all attempts, we still need to improve our indexing systems. We should have at least one well-designed and up-to-date web-based database, indexing Iranian scientific articles. Such a task will definitely need to be supported by the government.

We suggest that MOHME and MOSRT enforce a set of rules designed to facilitate indexing. Each registered journal must be required to follow these rules in order to extend its registration. According to the shortcomings we observed in this study, we suggest the information listed in Table 8 be provided by all registered journals for each issue published. The data should be sent electronically to the ministries before paper publication (using XML). The ministries should add an ID to each paper (similar to the PMID of PubMed) and make these data available to the indexing sites. Each journal should provide its abbreviated name to the ministry and Persian-language journals should provide their names in English. These data should be made available to registered journals and in turn to authors. Journals should require authors to use these names in their references instead of arbitrary translations of Persian journal names. The journal administration should be held responsible for errors in references and their format. The compliance of journals to these rules should be included in their rankings.

A good indexing system should have an English interface and include all the information listed in Table 8. Searches must be in English and a single search should retrieve both English and Persian-language papers. Persian search should be limited to author names. All search options listed in Table 4 should be available. Access should be free and a good online help should be available. Full text is not required but links to the journal web site or full text should be present when available. If the information is received electronically, as suggested above, the maintenance of such a site would be much less expensive and could be easily managed by a governmental grant. The site will be able to provide up-to-date information in addition to a reliable citation index. We suggest that only new issues be entered in to the system and past issues should be added gradually.

We believe that if our suggestions are followed, an international-class indexing system including all registered scientific journals published in Iran is within reach in one or two years.

Acknowledgment

The authors wish to appreciate the contributions of Dr. Seyed-Mehrdad Mohammadi and Dr. Reza Yousefi-Nourayee for their comments on methodology. The study was funded by a grant from Iran National Science Foundation.

References

1 Radmard AR, Khademi H, Azarmina P, Sadat-Safavi M, Nouraie M, Kolahdouzan S, et al. Iran's biomedical sciences' research output in 2003: a bibliographic analysis of medline and Excerpta Medica databases. *Arch Iran Med*. 2005; **8:** 180 – 183.

2 Moin M, Mahmoudi M, Rezaei N. Scientific output of

Iran at the threshold of the 21st century. *Scientometrics*. 2005; **62:** 239 – 248.

3 Malekzadeh R, Mokri A, Azarmina P. Medical science and research in Iran. *Arch Iran Med.* 2001; **4:** 27 – 39.

4 Butler D. Islam and Science: the data gap. *Nature*. 2006; **444:** 26 – 27.

5 Abbassi Z. Criteria for evaluating informa*tion b*anks [in Persian]. *IranDoc eJournal*. 2006;6. Available from: URL: http://www.irandoc.ac.ir/data/E_J/vol6/abbasi.htm

6 Curtis D. Evaluation criteria for online databases. Available from: URL: http://www.bcps.org/offices/lis/ila/rubrics/Evaluation%20Criteria%20for%20Online%20Databases.doc