DESIDOC Journal of Library & Information Technology, Vol. 33, No. 4, July 2013, pp. 343-351

DRDO Knowledge Repository

Anil Kumar Singh and Sumati Sharma

Defence Scientific Information and Documentation Centre (DESIDOC), Delhi-100 054
E-mail: aksinghto@gmail.com, sumati_s@yahoo.com

ABSTRACT

This paper deals with knowledge management and its need and benefit to R&D organisation. It describe the importance of standard format or guidelines for writing and assigning uniform numbers to technical reports in DRDO. It Ddiscusses in detail the establishment of DRDO Knowledge Repository (KR). Main emphasis has been given on how to use DSpace as open source software, scope of work in DRDO KR, installation and customisation of input data field, metadata creation, and uploading.

Keywords: Knowledge management, DSpace, knowledge repository, report writing

1. INTRODUCTION

The main objective of any R&D organisation is to conduct research in its area of work and design and develop products based on the demand of the users. People involved in the R&D gaine experiences and their tactic knowledge also gets enriched. So, main driving force in any R&D organisation is people and their knowledge. Organisations need to utilise this knowledge in the most efficient way because it is a fundamental resource that allows us to function intelligently. On the other side, the tremendous growth of information and communication technologies has made the organisations to continuously strive to improve their knowledge and overall performance. After the information explosion through internet, organisations have become more conscious towards information management and now they are evolving as knowledge organisations.

As knowledge becomes the key strategic asset for organisations, they start thinking to invest in creating, storing, and implementing the best knowledge networks, processes, methods, tools and technologies in the form of knowledge management. This will enable them to learn, create new knowledge, and apply the best knowledge much faster. The complete knowledge management in any R&D organisation is a systematic process by which knowledge is created, captured, stored, shared, and leveraged as follows:

- Creation: Knowledge is created in R&D organisation by gathering some data through experimentation, any R&D work, improving existing work, running day to day work, etc. The R&D in organisation generates millions of data/information.
- Capture: Creators of knowledge are not aware of this and valuable information/knowledge may therefore be lost. To prevent this information/knowledge, one needs to capture this knowledge.
- Storage: Once knowledge is captured, organisation and its people are aware of its existence. If the captured knowledge is relevant then it has to be stored, so that it can be used by those who need it in the organisation.
- Sharing and Application: After knowledge is properly documented and stored, it's needed to be shared and applied. This is the most important task in the knowledge management. It is not enough to install and fill a knowledge repository (KR) and expect the organisation to suddenly start sharing knowledge. Particular attention should be paid to building a knowledge sharing culture in the organisation.
- Technology: Technology can only play a supporting role in knowledge sharing and application. It can make knowledge sharing more effective and easier. Intranets have emerged as one of the most effective ways of sharing information and knowledge within R&D organisation.
- Innovation: Successful as well as effective knowledge sharing and application also stimulates innovation, improvement in existing knowledge and creation of new knowledge. This essentially closes the knowledge cycle.

Received 21 November 2012, online published 22 July 2013

2. NEED AND BENEFIT OF KNOWLEDGE REPOSITORY

The R&D organisation house treasures of knowledge that have been hidden in archives, almirahs, print formats, and a variety of storage media. These treasures include scientific, technological, cultural, artistic and historical materials generally unavailable to searchers and the public. This makes access to these information/ knowledge content very difficult. The organisation ends up having tons of information/ knowledge content stored but not used. Organisations need to gather and organise all the information that they need to know for their survival and then pool the captured information at a common place as a repository. This type of repository is often known as institutional or knowledge repository which offers solutions to such type of problems. Such an organisational information repository becomes knowledge repository when it is analysed, adapted and used effectively for the organisational end.

Having readily accessible, well organised centrally located digital information in the form of KR can be a tremendous asset for any R&D organisation. The primary benefit of knowledge or institutional repository is that it will enhance intellectual quality, reputation and visibility of the organisation. Institutional manager can develop procedure for using the repository to answer question such as:

- What result institution gets for the money which was put for R&D projects?
- What is being published from this institution?
- Who is producing what?
- Where is the upward and downward trend?
- How much collaborative work is being done and with whom?

In nutshell one can say that KR benefits are for:

- Contributors, because their articles are cited more
- Organisation, because documents produced by the organisation is available at one place, reflecting the intellectual achievement of the organisation
- Users, because documents in an institutional repository can be found through a search engine.

Many organisations have realised the value of institutional knowledge and already initiated knowledge management systems that collect, store, redistribute, utilise and ultimately leverage the institutional knowledge for the benefit of the organisation[1]. R&D institution like DRDO also felt the need to develop knowledge repository. Keeping this in view, DESIDOC has planned to create DRDO Knowledge Repository and has presently undertaken this pilot project. This paper discusses some of the details of the initiative taken by the Knowledge Management Division of the DESIDOC.

3. DRDO KNOWLEDGE REPOSITORY

DRDO Knowledge Repository is a computerised system that systematically collects, digitise, preserve, and disseminate in digital form, the intellectual output of DRDO. The focus of this knowledge repository is on collecting, digitising, organising knowledge generated by the DRDO laboratories across India in the form of various types of R&D reports. R&D reports are the primary and vital source of information about current or projected research. These reports contain valuable information and form the intellectual capital of DRDO which helps in achieving organisation's goal and also acts as a catalyst for organisation change.

3.1 Need and Purpose

In an increasingly competitive and volatile world, the ability to share knowledge is a prerequisite for successful growth, especially for organisation like DRDO that prize intellectual capital as one of their most valuable assets. Today DRDO with 52 laboratories across India is engaged in developing defence technologies covering various disciplines like aeronautics, armaments, electronics, combat vehicles, naval systems, life sciences, training, information systems, and agriculture. A number of projects are taken up by these laboratories for carrying out R&D in these areas. As a result of this, a large number of R&D reports are generated by these laboratories. Apart from DRDO laboratories, the Technical Directorates, and Research Boards of DRDO also generates considerable number of reports and publications through sponsored research. A variety of reports like design report, project report, technical report, feasibility report, progress report, field test report, project proposal report, validation report, trial report, technical memorandum, review report, research report, technical specification, interim report, annual reports, special publications, etc. are brought out and are distributed to limited extent.

Earlier there was no centralised mechanism to collate, index, provide access to and retrieve a report when needed at a later date. Also, no standard format was followed and it was not uncommon to find different formats and styles in reports brought by different Projects/Directorates or Divisions in the same laboratory. Further, there was no unique report number, as in the case of NTIS reports, to locate and retrieve the report when needed.

Keeping this in view, Defence Research Council tasked DESIDOC to devise a mechanism to collect, collate, index, and digitise the unclassified/restricted

reports and provide access to them on DRDO intranet. The main objective was to strengthen the knowledge-base of DRDO with uniform style and procedures comparable to the international programs for report documentation like NTIS, NASA, RAND, etc. Accordingly, it was strongly proposed to create a DRDO Knowledge Repository of technical reports and to maintain a copy each of the unclassified and restricted reports at DESIDOC.

4. GUIDELINES FOR WRITING AND ASSIGNING UNIFORM NUMBERS TO TECHNICAL REPORTS IN DRDO

In DRDO, there was a need to have standard procedure or specification for writing reports and assigning appropriate report numbers for various types of reports generated in DRDO laboratories[2]. It could reduce the variation in the format, style of presentation, design, treatment, and categorisation in reports being brought out by large laboratories and where a number of projects are running, each one of them brings out its reports with varying styles and formats. This led to non-uniformity, lack of identity, as well as difficulties in timely retrieval of required reports. DESIDOC entailed to design a suitable numbering and control mechanism for the technical reports in accordance with the Indian Standard.

This guideline includes the following:

- Technical reports and their types
- Writing of technical reports
- Assigning standard numbers for DRDO technical reports with examples

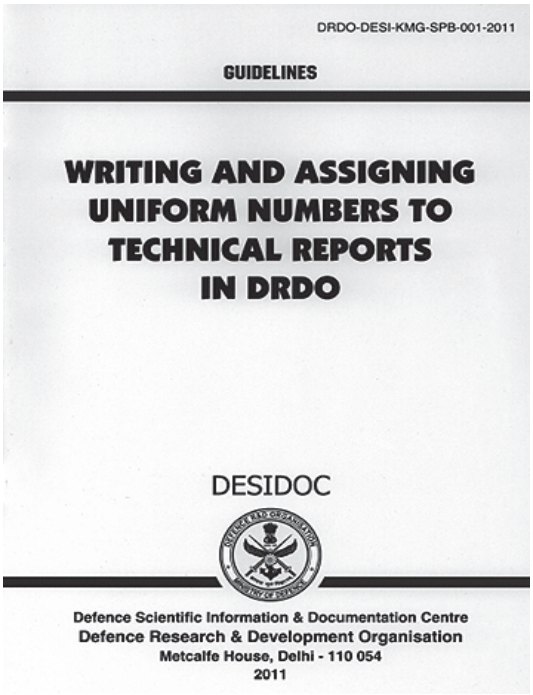


Figure 1. Guidelines for writing uniform numbers.

- Guidelines for completion of bibliographic description sheet, and
- Points to be considered while preparing a technical report

The purpose of this guideline (Fig. 1) is to provide a uniform format for writing technical reports and to enable DRDO Laboratories to assign uniform report number to the reports generated in their laboratory[2]. This would lead to compatibility in the report numbers assigned by the DRDO Laboratories so that the Technical Information Resource Centers and other technical report users are able to identify, locate, and organise report literature according to a consistent and accepted pattern.

5. ESTABLISHMENT OF DRDO KNOWLEDGE REPOSITORY

While establishing any repository, institution has to make lot of decisions. Policies, system architecture, and other elements will depend on institutional context and the scope and purpose of the repository. Some of the important key issues considered while developing DRDO Knowledge Repository are:

5.1 Organisation Culture

Organisation culture depends on how the organisation is structured. DRDO HQRs is situated in New Delhi and 52 laboratories are situated across India. Coordinating with all the laboratories was a tough task. It was decided that with a central repository at DESIDOC, all the laboratories need to be convinced that contributing to a central repository will enhance their reputation in their disciplines and result in wider dissemination of their work.

5.2 Scope

Repository must decide early on purpose and scope and communicate them to all the laboratories. It was decided to have four digitisation centres to cover all DRDO Labs across India viz. Delhi, Bengaluru, Pune, Hyderabad region.

5.3 Content

Decision making on content can become a tedious issue. Criteria for deposit into the repository could come from each community or from a central body with input from the users. Type of documents to be included in the repository is to be decided. DRDO Knowledge Repository includes all types of unclassified and restricted R&D reports published from DRDO laboratories.

5.4 Access

Accessibility of repository on internet or intranet is to be decided early on. Knowledge created

by DRDO in the form of R&D reports must be shared on DRDO intranet server. Accordingly it was decided to upload DRDO KR database on DRONA (DRDO Rapid On-line Network Access). Submission and access rights have been given to all DRDO laboratories and DRDO HQrs. Administration power remains with the DESIDOC.

5.5 Standards and Software

Institutional repository employ standards developed to handle issues associated with open access. These standards include OAIS (Open Archival Information Systems), OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting), METS (Metadata Encoding and Transmission Standards).

Software is the key element in the construction of any repository. There are number of software options available when setting up an institutional repository e.g., ePrints, DSpace, Fedora, Greenstone Digital Library Software, etc. These are all open source softwares. Some proprietary softwares are also available in the market. There are pros and cons to both proprietary and open source platforms. Although setting up the repository is fairly quick and easy on a proprietary platform, it can be very expensive to pay annual fees for hosting and using the service. For the open source software, it may appear to be 'free', however, installing and customising the repository can be very time consuming and these solutions often lack technical and development support. There is no uniform rule for choosing a platform. It depends on what the organisation wants to achieve.

5.6 Legal Aspects

Organisation responsible for operating and maintaining repository need to ensure that all legal requirements are met. These requirements include appropriate software, content licenses, and intellectual property rights.

5.7 Maintenance

Maintenance is the key issue that involves long-term commitment by the organisation. A repository cannot run by itself. It needs constant attention, maintenance of contents, backups etc. Staff in maintenance needs to know the consequences of changes in hardware, software, and standards and be able to adjust accordingly.

6. USE OF DSPACE AS CHOSEN SOFTWARE

For DRDO KR DSpace software was identified, installed and tested because of several reasons[3]:

- DSpace is a Java-based system developed by MIT and HP labs
- Uses highly scalable Postgre SQL as the database back-end
- It offers a communities-collections model and has built-in submission workflows and long-term preservation function
- Compliance with standards and protocols like DC, OAI-PMH, UNICODE, etc.
- It uses Lucene search engine which offers decent search capability
- Intuitive way of self archiving. Capable of archiving digital items in various formats
- Flexible for data migration
- Uses widely accepted metadata scheme
- Offer customisation facility which did not involve much programming
- Majority of the repositories worldwide are created using the DSpace platform

6.1 Installation of DSpace

After taking the decision to use DSpace , the first challenge faced was the installation. DSpace runs on Postgre SQL and requires server installation. Although DSpace runs on a variety of operating systems it was decided to purchase Red Hat enterprise Linux after some testing because it is the recommended OS for DSpace. During installation some difficulties were noticed, later on it was rectified and DSpace was successfully installed[4].

6.2 Customisation of DSpace

To get the desired customisation according to need of DRDO KR necessary changes were made in DSpace by using DSpace.Cfg and trying other facilities like Manakin etc. Manakin (XMLUI) is a web-based, XML driven user interface to DSpace that introduces modular interface layer, enabling an institution to easily customise the interface according to the specific needs of the particular repository[5]. Home-page of DRDO KR was designed (as shown in Fig. 2). Two communities were created (as shown in Fig. 3):

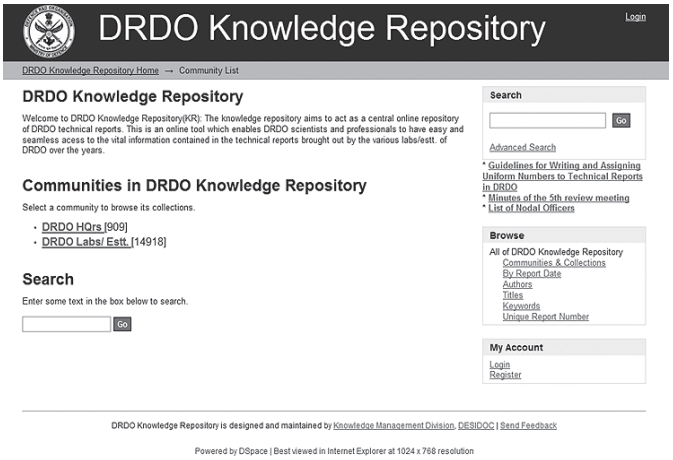


Figure 2. Home page of DRDO Knowledge Repository.

(a) DRDO HQrs

(b) DRDO Labs/Estt

Under DRDO HQrs three sub communities were created:

- Defence Science Organisation
- Research Boards
- Technical Directorates

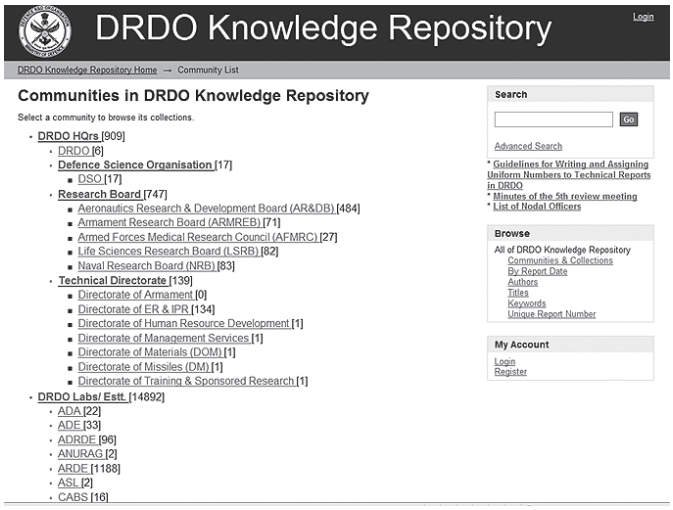


Figure 3. Communities in DRDO Knowledge Repository.

6.3 Customisation of Input Data Fields

DSpace uses predefined Dubiln Core standard metadata format. Metadata input form is used while submitting an item to the collection. This metadata input form was customised according to the need of DRDO technical reports (Fig. 4). These reports contains fields like author, title, year, performing lab, unique report no., report type, subject, abstract, security classification, etc. The customisation of metadata input form is necessary to add some of the additional metadata to the items. Additional metadata added to DRDO technical reports were unique report no., performing lab, sponsor, report type, security classification etc.

In DSpace, metadata entry forms are controlled by a single XML file, input-forms.xml, in the config subdirectory under DSpace home. DSpace comes with a sample configuration that implements the traditional metadata-entry forms[5]. The customisation options of metadata input form can be:

- The number of metadata-entry pages
- Which field appear on each page, and their sequence
- Labels, prompts, and other text associated with each field
- List of available choices for each menu-driven field

Whenever any changes made to the input-forms. xml file, one should always restart Tomcat (or other servlet container used) to take effect.

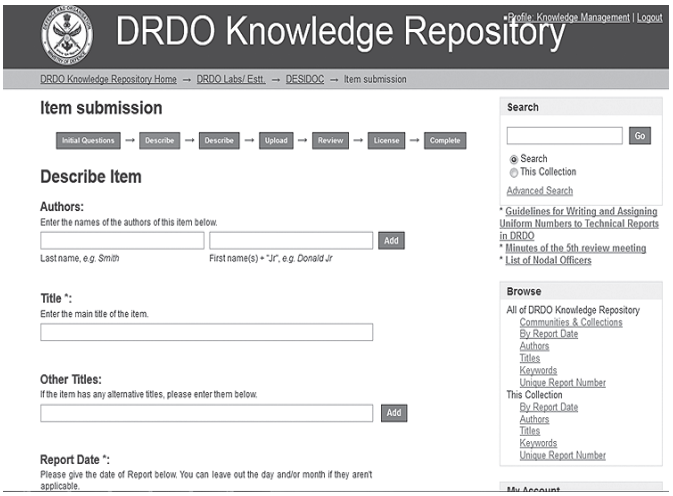


Figure 4. Customisation of metadata input sheet.

6.4 Standardisation of Data using Authority Files

Data inputted in metadata entry forms were standardised wherever it was possible and necessary[5]. Rendering of data in the fields like author, title, corporate names, sponsor, etc. were standardised by using AACR2 code. Authority file for personal names was maintained. Defense Technical Information Center, USA (DTIC) thesaurus was used as a controlled vocabulary tool for assigning subject/keywords. Since most of the technical terms in DRDO R&D reports were defence and military related, it was decided to use DTIC thesaurus[6]. Standardisation of data in any repository ensures quality, easy data exchange ,and interoperability.

7. SCOPE OF WORK IN DRDO KNOWLEDGE REPOSITORY

DRDO KR project involves four major tasks, viz.

(a) Collection of reports

(b) Unique report number assignments

(c) Scanning and digitisation, and

(d) Metadata creation, and uploading in DSpace.

7.1 Collection of Reports

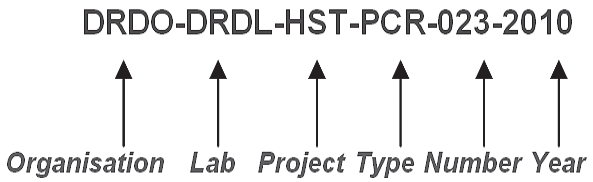
Technical reports were collected from DRDO labs situated across India as well as from Technical Directorates and Research Boards at DRDO HQrs, New Delhi.

7.2 Unique Report Number Assignment

Reports collected from Labs were assigned unique report number with the help of 'Guidelines for writing and assigning uniform numbers to technical reports in DRDO'[2]. Unique report number consists of the following elements:

1. Name of the parent organisation (DRDO),

2. Name of the laboratory (Maximum four alphabetical character),
3. Name of the project/programme/directorate/ division (Only three alphabetical character),
4. Type of report (Maximum of three character),
5. Report number (Continuous number given by Lab in numeric character), and
6. Year of publication (In four numeric character). For example



7.3 Scanning and Digitisation

Digitisation is the process of converting information into a digital format. It includes acquiring, converting, storing and retaining information in standardised and organised manner with technological support. The documents are being scanned in the resolution of 250 to 400 dpi (dot per inch) depending on the physical condition of the documents. After scanning, next process involves editing the scanned images for any error and makes text files error free. Photoshop is used to remove any noise, dust, shadow and other artifacts from scanned pages. Scan Tailor software is being used to straighten, crop and reorient the cleaned up pages. Scanned images are then processed for OCR using Acrobat's 'OCR text recognition option' and finally these are saved in the computer in PDF A formats with some file name.

7.4 Metadata Creation and Uploading in DSpace

DSpace software serves as a repository and stores digital content[5]. In a system with such a goal, perhaps the most critical aspect of the system is how that data enters the system. When users log in, documents which have been digitised as PDF are copied into the system where DSpace is installed and then they go through a configurable workflow where they describe and upload their submissions to the live repository in the following way:

Metadata is recorded in Dublin Core format where four fields are treated as mandatory e.g. Unique Report No., Title, Year and Performing Lab along with other available metadata as follows (Table 1):

Table 1. Metadata description for DRDO KR

S. No.	Submission workflow step	Description
1.	Describe	Describe the item if it has more than one title, more than one file
2.	Describe	Enter metadata about the document (s) for the following fields: author, title, report date, performing lab, unique report no., original report no., report type, language
3.	Describe	Assign subject/keywords, write abstract, mention sponsor if any, security classification
4.	Upload	Enter the name of the file to be uploaded
5.	Verify	Verify for any correction in the metadata fields and file uploaded
6.	License	Read the terms of the license or agreement. If ok click 'Grant License'
7.	Complete	Submission is completed

Unique Report Number

DC Element:	relation
DC Qualifier:	unique report number;
Requirement:	Mandatory;
Occurrence:	Once only;
Format:	DRDO-[Performing lab name]-[type of report/original report no.]-[yyyy];
Example:	DRDO-DIPA-PCR-001-1999;

Author Name

DC Element:	contributor
DC Qualifier:	author;
Requirement:	Optional;
Occurrence:	Max. Three names only;
Format:	[Last name, First name], [Last name, First name], [Last name, First name];
Example:	Mathur, A., Ganguli, C.;

Report Title

DC Element:	title;
DC Qualifier:	none;
Requirement:	Mandatory;
Occurrence:	Once only;
Format:	Text; First letter of the title must be in CAPS, rest in small; Acronym and abbreviation should be captured as ALL CAPS; First letter of nouns should be captured as CAPS;
Example:	Guidelines for writing and assigning uniform numbers to technical reports in DRDO;

Other Title

DC Element:	title;
DC Qualifier:	alternative;
Requirement:	Optional;
Occurrence:	Once onLy;
Format:	Text; First letter of the titlle must be in CAPS, rest in small; Acronym and abbreviation should be captured as ALL CAPS; First letter of nouns should be captured as CAPS;
Example:	User handbook

Report Date

DC Element:	date;
DC Qualifier:	reportdate;
Requirement:	Mandatory;
Occurrence:	Once only;
Format:	[yyyy];
Example:	1999;

Performing Lab/Inst

DC Element:	identifier;
DC Qualifier:	performinglab;
Requirement:	Optional;
Occurrence:	Once only;
Format:	[ACRONYM, Fixed Values Only, LENGTH NOT FIXED];
Example:	DIPA;

Pages

DC Element:	identifier;
DC Qualifier:	pages;
Requirement:	Optional;
Occurrence:	Once only;
Format:	[n/nn/nnn, LENGTH NOT FIXED];
Example:	45;

Original Report Number

DC Element:	relation;
DC Qualifier:	originalreportnumber;
Requirement:	Optional;
Occurrence:	Once only;
Format:	[Misc, LENGTH NOT FIXED];
Example:	DRDO-12321;

Language

DC Element:	language;
DC Qualifier:	iso;
Requirement:	Optional;
Occurrence:	Once only;
Format:	[Fixed Values Only, LENGTH NOT FIXED];
Example:	English (ISO Code of English;

Report Type

DC Element:	type;
DC Qualifier:	reporttype;
Requirement:	Optional;
Occurrence:	Once only;
Format:	Fixed Values Only, LENGTH NOT FIXED];
Example:	Acceptance Test Procedure-ATP;

Keywords

DC Element:	subject;
DC Qualifier:	none;
Requirement:	Optional;
Occurrence:	Multiple (Unbound);
Format:	[Misc];
Example:	Armament, Pinaka weapon,Rockets,Trial report;

Abstract

Element:	description
Qualifier:	abstract;
Requirement:	Optional;
Occurrence:	Once onLy;
Format:	[Misc, LENGTH NOT FIXED];
Example:	This paper examines the evolution of the U.S. interest swap market. The authors review theory and past empirical studies on U.S. swap spreads and estimate an error-correction model for maturities of 2, 5, and lOycars from 1994 to 2004;

Sponsorship

DC Element:	description;
DC Qualifier:	sponsorship;
Requirement:	Optional;
Occurrence:	Once only;
Format:	[ACRONYM, Fixed Values Only, LENGTH NOT FIXED;
Example:	ADE;

Classification

DC Element:	subject;
DC Qualifier:	classification;
Requirement:	Optional;
Occurrence:	Once onLy;
Format:	[Fixed Values Only, LENGTH NOT FIXED];
Example:	Unclassified;

8. DISSEMINATION

The document submitted and archived into the DRDO Knowledge Repository (KR) can be disseminated and accessed by the users through search and browse (as shown in Fig. 6 & 7). Repository offers users the capability to search the data base for items of interest in both simple and advanced form. From the DRDO KR home-page, users can browse all items by

- Communities and collection
- Report date

- Author
- Title
- Subject
- Unique report number

9. CONCLUSIONS

DRDO Knowledge repository (KR) is developed considering the advantages like no physical boundaries, round-the-clock availability of DRDO technical reports, multiple accesses to information resources, faster information search and retrieval, preservation and conservation of exact copy of the original document. Access and retrieval has been increased in all the DRDO labs as well as DRDO HQrs.

Although DRDO Knowledge Repository(KR) was successfully built and implemented it was experienced that building a successful knowledge repository(KR) was much more beyond 'scanning' and 'putting document online'. In the process, though lot many challenges were faced like preparation of guidelines, collection of reports across India, digitisation issues, constant updating of records, regular maintenance and backups, DESIDOC was able to create a knowledge repository for the knowledge generated by the DRDO and could provide to all DRDO labs through its intranet DRONA from where scientists/staff of various labs are accessing this repository which at present contains more than 15760 records.

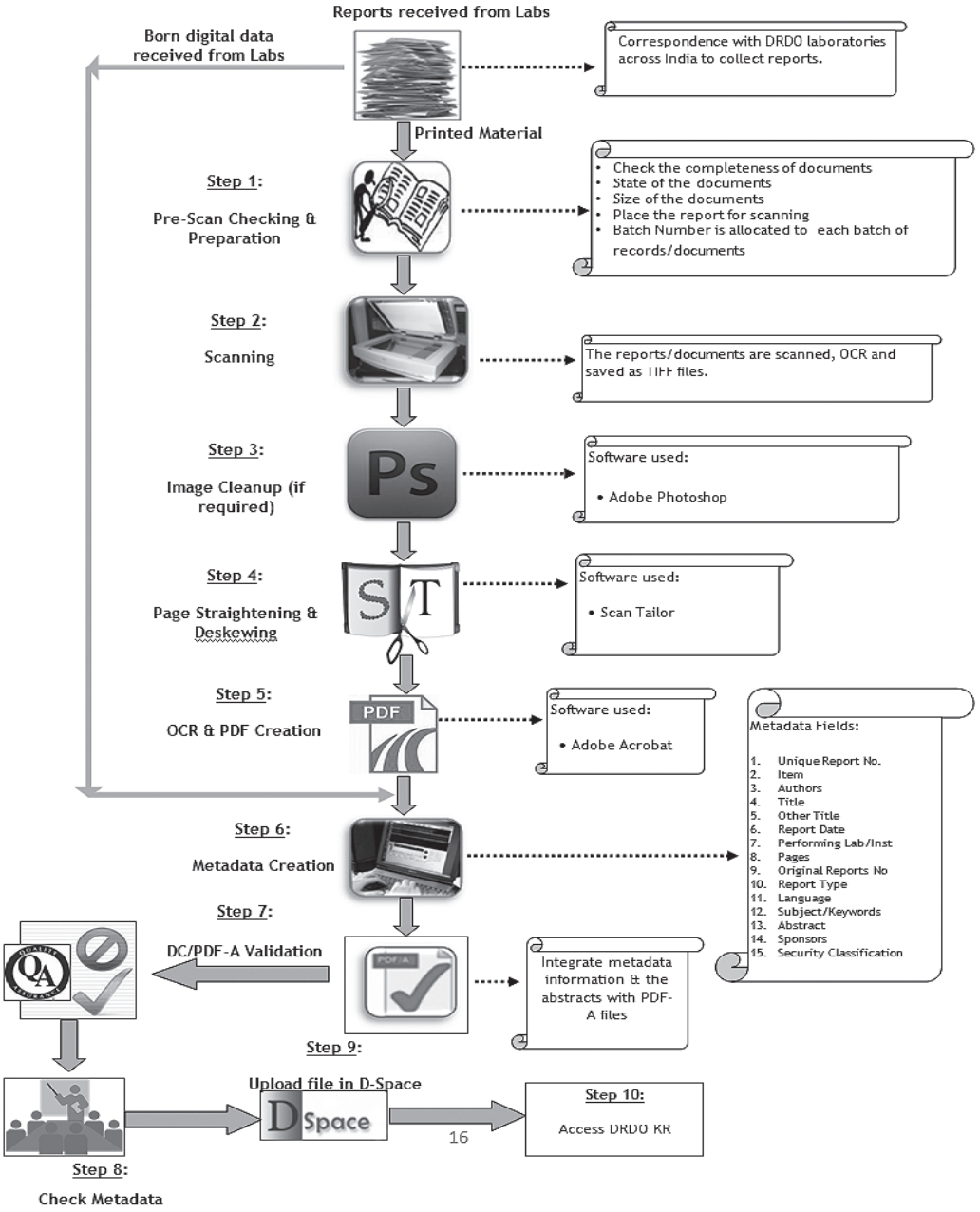


Figure 5. DRDO KR workflow.

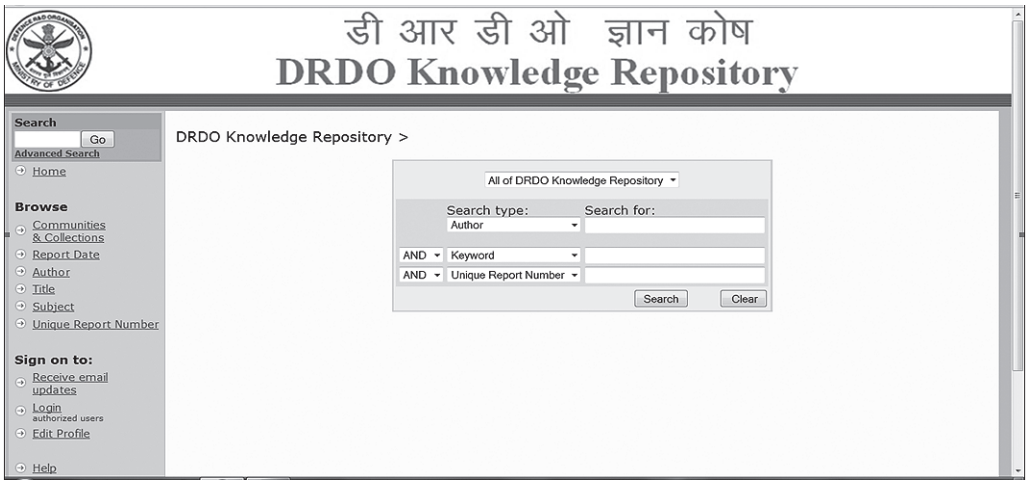


Figure 6. Advance search in DRDO KR.

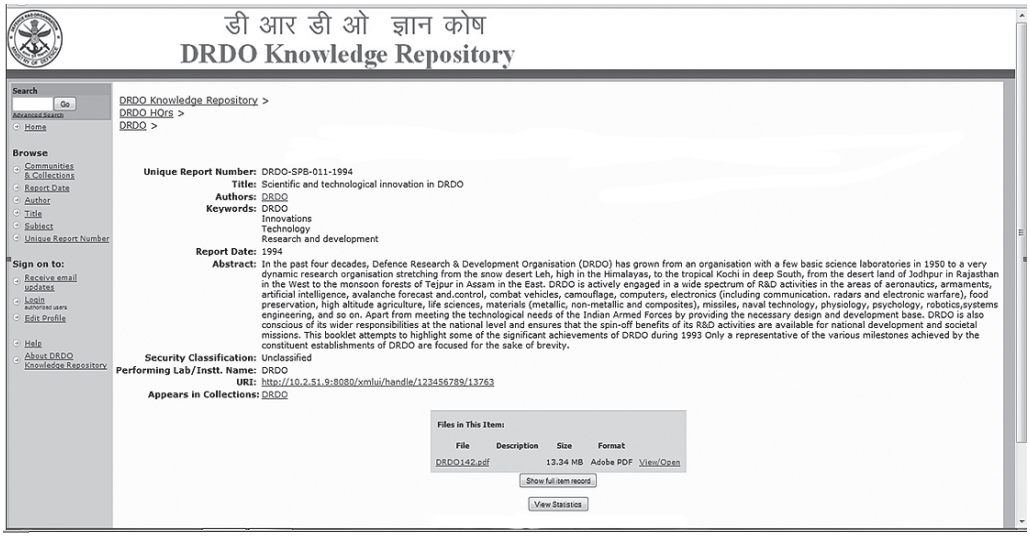


Figure 7. Sample search result in DRDO KR.

REFERENCES

1. Ashalatha, *et al.* Building institutional repository: A overview.*OCLC Sys. & Serv.*, 2007, **23**, 278-86.
2. Guidelines for writing and assigning uniform numbers to technical reports in DRDO. DESIDOC. Metcalfe House, Delhi.
3. Thakuria, Juli. Building an institutional repository with DSpace. 6th Convention Planner-2008, Nagaland University, Nagaland, 6-7 November 2008.
4. DSpace. http://www.dspace.org/.
5. Prasad, A.R.D. & Madalli, Devika P. Advanced DSpace. *In* DRTC Workshop 2011, 17-18 February 2011, Documentation Research and Training Centre, Bangalore.
6. DTIC. http://www.dtic.mil/dtic/stresources/techreports/dticSearchTools/thesaurus_desc.html/.