

SITED User's Guide and Data Management Structure

Author: Kristopher L. Kuhlman
(6212) Repository Performance Department

1	User's Guide for SITED Platform.....	1
1.1	Login to the SITED webpage	2
1.2	SITED functions on existing records.....	4
1.2.1	Browsing SITED existing records.....	4
1.2.2	Searching SITED existing records	4
1.2.3	SITED search results page.....	5
1.2.4	Edit existing records.....	6
1.3	SITED functions to add records.....	7
1.3.1	Add a single record to SITED	7
1.3.2	Import multiple records to SITED	8
1.4	Errors	8
2	Data Management Structure	9
2.1	Database Deliverables for WIPP Salt Data Compilation and Assessment.....	9
2.2	Bibliographic Database Design	9
2.2.1	Bibliographic Database Layout.....	9
2.2.2	SRDI Keywords.....	10
2.3	Initial population of database	11
2.4	Electronic copies of available data.....	11

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

1 User's Guide for SITED Platform

The Salt Investigations Technical Expansive Database (SITED) is a bibliographic database based upon the Refbase open-source project. The records are stored in a MySQL database, the web interface is generated using PHP, and served up via the Apache webserver. The web interface to SITED is <https://sited.sandia.gov/sited>, which is externally visible and password protected. The SITED web interface facilitates the primary tasks:

1. Adding new database entries;
2. Searching (querying) existing database entries;
3. Bulk importing bibliographic references in several common formats.

The Refbase software (<http://www.refbase.net>) has documentation (<http://www.refbase.net/index.php/Documentation>) and Frequently Asked Questions (FAQs) on the refbase website, but we describe the main tasks required by most database users.

1.1 Login to the SITED webpage

Navigate to <https://sited.sandia.gov/sited> and log into the database (see red highlighted area in Figure 1).

If you receive a warning or error about the certificate of the website being invalid, please continue. In Internet Explorer choose “continue to website (not recommended).” In Firefox choose “I understand the risks > allow exception.” This error only exists as the server transitions to its final configuration.

If you do not have an account, contact the administrator to obtain one (Kris Kuhlman k1kuhl1m@sandia.gov, (505) 845-0938).

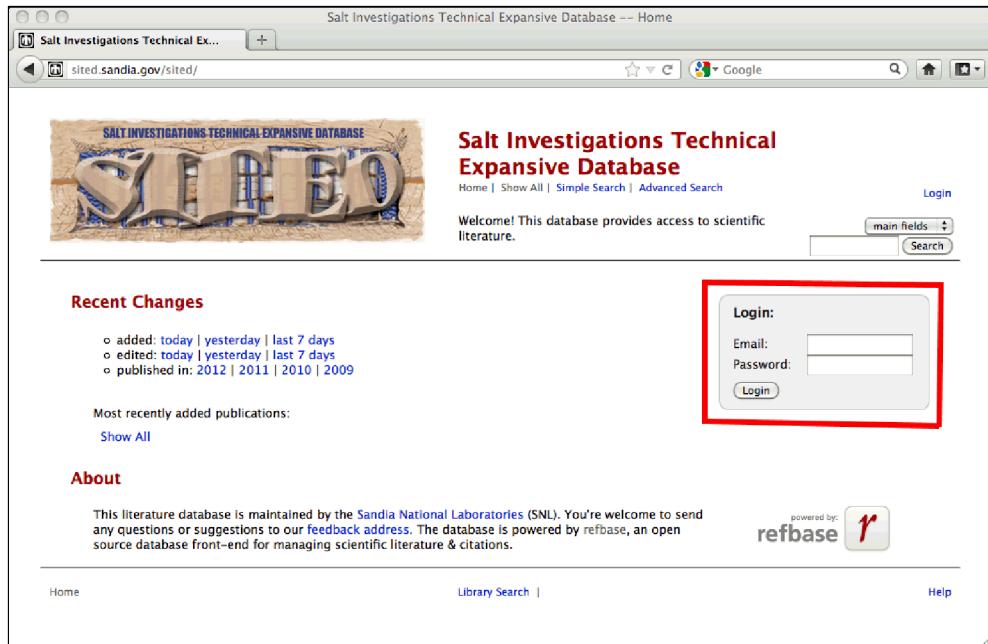


Figure 1. SITED login box highlighted in red

Once you have logged in, your username will appear in the upper right corner, and the main list of options is given below the title in red (see Figure 2):

“Home | Show All | Simple Search | Advanced Search | Add Record | Import”

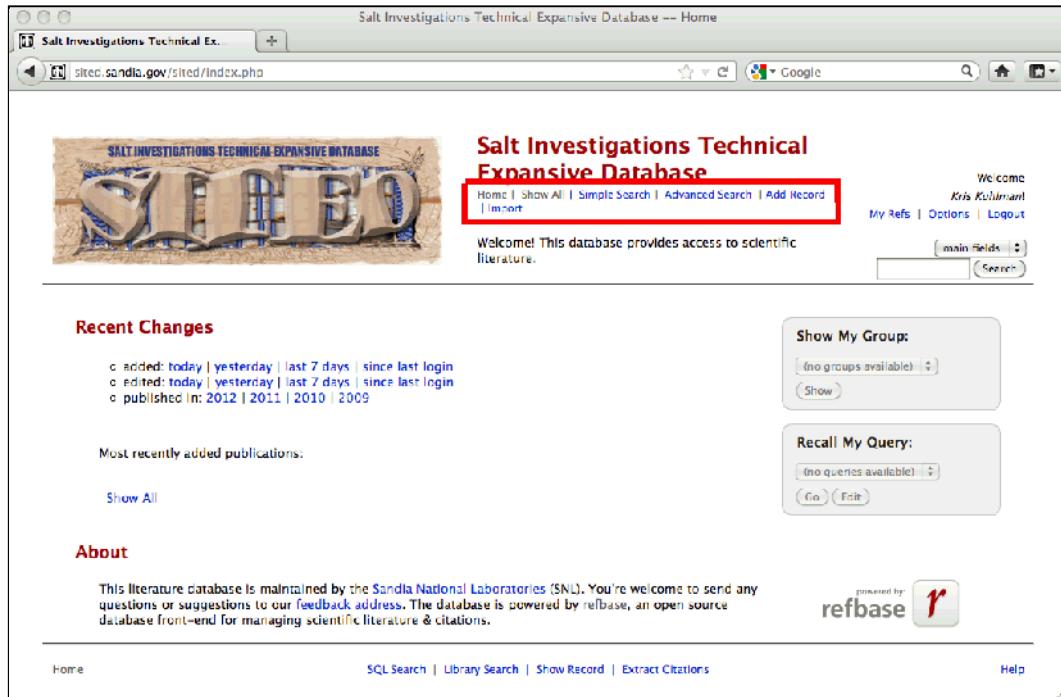


Figure 2. SITED main page after logging in. Main navigation options are highlighted in red.

1.2 SITED functions on existing records

1.2.1 Browsing SITED existing records

The “Show All” option on the menu lists all records in the database.

1.2.2 Searching SITED existing records

There are three searching options. “Simple Search” and “Advanced Search” are similar, except advanced search presents more options. The third option is the search box and search button in the upper right of the main screen (a.k.a. “Simplest Search”). For many purposes, such as searching for a keyword or author, this simplest search is best. The simplest search automatically queries the “Author”, “Title”, “Publication”, “Keywords”, and “Abstract” fields (but is configurable using the pull-down menu). Simple search includes the “Author”, “Title”, “Publication”, “Volumes”, and “Pages” fields.

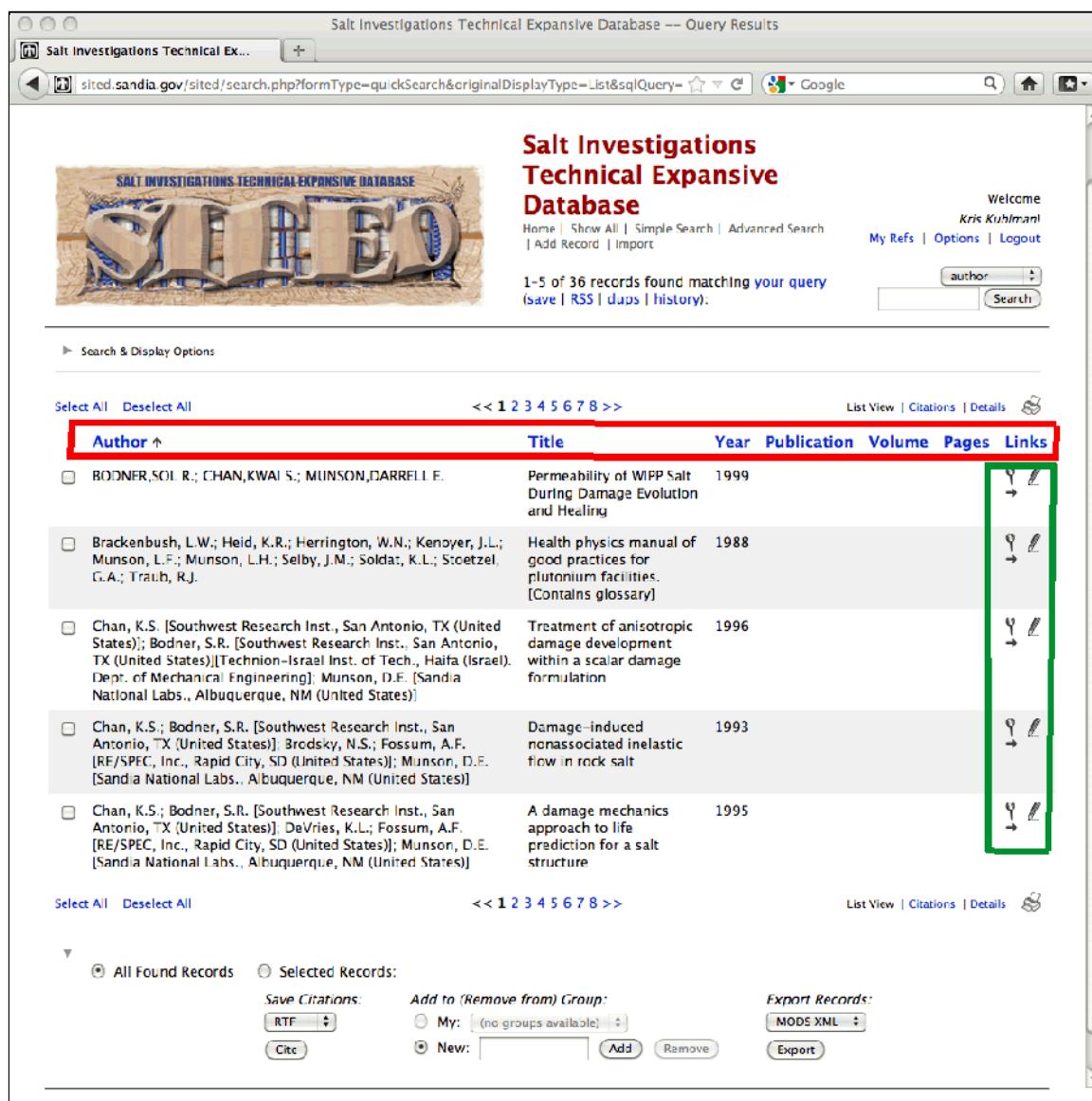
All search results lead to a similar results page, described next. The browsing results (see section 1.2.1 above) lead to the same results as a search for the empty string, which every entry matches.

1.2.3 SITED search results page

An example query was performed for the author “Munson” (see Figure 3). Five results are shown per page by default. The user can navigate through pages with the numbers “<< 1 2 3 4 5 6 7 8 >>” above the column headings.

The results are default sorted alphabetically by “Author” (see the up arrow  by the heading “Author” in the red highlighted area of Figure 3); clicking one of the other headings re-sorts the search results. Clicking the same heading changes the sort order (ascending or descending).

To view a record, click the magnifying glass icon at the right; to edit or delete a record, click the pencil icon at the right (both are highlighted in green in Figure 3).



The screenshot shows the SITED search results page for the query "Munson". The page title is "Salt Investigations Technical Expansive Database -- Query Results". The search bar shows the query "Munson". The results table has columns: Author, Title, Year, Publication, Volume, Pages, and Links. The "Author" column is highlighted with a red border. The "Links" column is highlighted with a green border. The results are as follows:

Author	Title	Year	Publication	Volume	Pages	Links
BODNER,SOL R.; CHAN,KWAI S.; MUNSON,DARRELL E.	Permeability of WIPP Salt During Damage Evolution and Healing	1999				 
Brackenbush, L.W.; Heid, K.R.; Herrington, W.N.; Kenoyer, J.L.; Munson, L.F.; Munson, L.H.; Selby, J.M.; Soldat, K.L.; Stoetzel, G.A.; Traub, R.J.	Health physics manual of good practices for plutonium facilities. [Contains glossary]	1988				 
Chan, K.S. [Southwest Research Inst., San Antonio, TX (United States)]; Bodner, S.R. [Southwest Research Inst., San Antonio, TX (United States)][Technion-Israel Inst. of Tech., Haifa (Israel)]. Dept. of Mechanical Engineering; Munson, D.E. [Sandia National Labs., Albuquerque, NM (United States)]	Treatment of anisotropic damage development within a scalar damage formulation	1996				 
Chan, K.S.; Bodner, S.R. [Southwest Research Inst., San Antonio, TX (United States)]. Brodsky, N.S.; Fossum, A.F. [RE/SPEC, Inc., Rapid City, SD (United States)]; Munson, D.E. [Sandia National Labs., Albuquerque, NM (United States)]	Damage-induced nonassociated inelastic flow in rock salt	1993				 
Chan, K.S.; Bodner, S.R. [Southwest Research Inst., San Antonio, TX (United States)]. DeVries, K.L.; Fossum, A.F. [RE/SPEC, Inc., Rapid City, SD (United States)]; Munson, D.E. [Sandia National Labs., Albuquerque, NM (United States)]	A damage mechanics approach to life prediction for a salt structure	1995				 

At the bottom, there are buttons for "Select All" and "Deselect All", a page navigation bar (<< 1 2 3 4 5 6 7 8 >>), and buttons for "List View", "Citations", and "Details".

Figure 3. Search results for "Munson" showing column headings (red) and read/edit actions (green)

At the bottom of the search results page, an exporting functionality is provided in several common formats.

1.2.4 Edit existing records

The database has been initially populated from existing databases, but there are fields that are specific to the Salt Research & Development Investigation (SRDI). We use the edit functionality to add this project-specific information as each record is reviewed. The same basic page is used for editing a record and creating a new record; when in editing mode the form is populated with the current entry, rather than the default blank state. Any changes made are saved and date-time and name of user that made the changes is recorded in the database. The editing page is described in the next section.

1.3 SITED functions to add records

There are two main ways to add records to SITED. They can be entered directly into the web interface, or they can be imported using the bulk import functionality.

1.3.1 Add a single record to SITED

The form to add a single record to SITED presents the user with a textbox or pull-down form for each valid field in the database (see Figure 4). The user should enter multiple entries in a single field (multiple authors, multiple keywords, etc.) separated by semicolons.

SALT INVESTIGATIONS TECHNICAL EXPANSIVE DATABASE
SITED

**Salt Investigations
Technical Expansive
Database**

Welcome
Kris Kuhlman!
My Refs | Options | Logout

Add a record to the database:

main fields

Author	<input type="text"/>	<input type="checkbox"/> is Editor
Title	<input type="text"/>	Type <input type="button" value="Report"/>
Year	<input type="text"/>	Publication <input type="text"/>
Volume	<input type="text"/>	Abbreviated Journal <input type="text"/>
Keywords	<input type="text"/>	Pages <input type="text"/>
Abstract <input type="text"/>		
Address <input type="text"/>		
Corporate Author <input type="text"/>		
Publisher	<input type="text"/>	Place of Publication <input type="text"/>
Language	<input type="text"/>	Editor <input type="text"/>
Original Title <input type="text"/>	Summary Language <input type="text"/>	
Abbreviated Series Title <input type="text"/>	Series Editor <input type="text"/>	
Edition <input type="text"/>	Series Volume <input type="text"/>	
Medium <input type="text"/>	ISSN <input type="text"/>	
Conference <input type="text"/>	Area <input type="text"/>	
Approved <input type="radio"/> yes <input checked="" type="radio"/> no	Notes <input type="text"/>	
Location your name & email address will be filled in automatically		
Call Number <input type="text"/>	Serial <input type="text"/>	not assigned yet
URL <input type="text"/>	DOI <input type="text"/>	
<input type="checkbox"/> Online publication. Cite with this text: <input type="text"/> <input type="checkbox"/> SNL publication		
<input type="button" value="SITED Descriptor"/> <input type="button" value="Rating"/> Archive if more than 5 uploads are needed. <input type="button" value="File Type"/> <input type="button" value="File Location"/>		

Figure 4. SITED entry page for adding a single record

1.3.2 Import multiple records to SITED

Use the bulk import functionality when multiple records are being imported from an external source. This maps the fields from several common formats (bibtex, endnote, etc.) to the format of the SITED database.

Records can either be imported from a text file (pick the file using the “browse” button), or the records can be copied and pasted directly into the textbox in the browser (see Figure 5).

Figure 5. SITED bulk import form

If additional importing functionality is required beyond that provided in the import page, contact the administrator (k1kuhl1m@sandia.gov). Some importing can be done to the database directly, either using SQL statements or the phpMyAdmin administration interface.

1.4 Errors

If errors or warnings are encountered during the use of the database, please contact the administrator (k1kuhl1m@sandia.gov) with a description of what you were doing or trying to do and the text of any error messages that appear in the browser.

2 Data Management Structure

2.1 Database Deliverables for WIPP Salt Data Compilation and Assessment

Discussed below in more detail, the work associated with Activity 1 will result in three main deliverables:

1. Salt Investigations Technical Expansive Database (SITED) is an electronic bibliographic database of SAND reports, studies, journal articles, and conference proceedings related to salt and repositories. SNL work will center on data and reports collected underground at WIPP, but the database will accommodate data from several other projects and locations.
 - a. A browser-based web interface to SITED to allow data entry and database querying from any location.
 - b. Standard bibliographic information on each reference: e.g., author(s), reference/call number, title, number of pages, abstract.
 - c. Electronic printable (Adobe PDF) versions of reports/documents stored on the server along with the database, linked from respective entries.
 - d. Electronic searchable (plain text) versions of reports.
2. A collection of electronically stored data for SRDI model validation and benchmarking activities. Data will be referenced from SITED. Data exist in a wide range of data types, file formats, and dataset sizes.
3. A summary report including:
 - a. A prose summary of SRDI-relevant work done at WIPP and description of the data available for use in other SRDI activities.
 - b. Description of the process used to construct SITED and curate the data contained in it.
 - c. Summary and statistics regarding the contents of SITED.

2.2 Bibliographic Database Design

SITED uses the MySQL open-source relational database. The browser-based interface is standard HTML created dynamically using PHP. Users can log into and use the database from any computer with an internet connection and a web browser, based upon the open-source Refbase database (<http://www.refbase.net>). We have used this existing open-source project to quickly create the required database and interface.

The database is developed using freely available software. The database is hosted on an SNL webserver with public visibility (<https://sited.sandia.gov/sited>), to facilitate the ability to view and edit data by personnel from multiple agencies (e.g., LANL, SNL Carlsbad, SNL Albuquerque, and external consultants).

2.2.1 Bibliographic Database Layout

The main table in SITED contains the following fields:

- common bibliographic information:
 - Author(s)
 - Title
 - Type (journal paper, report, conference proceedings)
 - Date
 - Number of pages
 - Abstract / Table of Contents
 - Conference Name or Journal Name (when applicable)
- path on server to electronic copy of report (PDF format)
- path to text copy of report (plain text format)
- SRDI-related metadata
 - Name of related QA Plan
 - Name of related Test Plan
 - Field to include if data or source code tabulated directly in the report?
(may be available via OCR)
 - Field to list path to data available electronically
 - SRDI Keywords (choose 0-5)
 - A few sentences summarizing or discussing the reference, especially related to its usefulness and applicability to other aspects of SRDI.

2.2.2 SRDI Keywords

SRDI Keywords will be chosen from the following table. Each keyword is assigned a default value of 0 (not applicable) to 5 (very applicable). Moving through the columns of the following table, one or more keywords (sometimes none) are chosen from each column to rank or locate the report in the conceptual hierarchy of salt research applicable to further activities in SRDI.

In general, the “5” level for an SRDI keyword is reserved for entries that are “required reading” for technical reasons, or are very important for historical reasons. Level “3” indicates a report is topical, but possibly not as important; it may be superseded by newer work. Level “1” signifies the report only mentions or is peripherally related to a topic, but might be useful for completeness reasons. Level “0” means not applicable (the default).

There is also a free-form “Keywords” field that is typically populated with a semicolon-delimited list of terms, originating from the source database (OSTI, SNL Technical Library, or WIPP Records Center). The SRDI Keywords are specific to the SRDI project and less flexible, to improve searchability. Free-form keywords often have poorer searchability due to trivial differences (salt vs. halite), mis-spellings, and inconsistent use between data sources.

Timing	Location/Scale	Report Class	Topic	Feature of Interest	Waste
pre-test	laboratory	summary	salt creep	anhydrite layers	DHLW
testing	in situ	data report	geophysics	argillaceous layers	TRU
post-test	modeling	test plan	heated salt	disturbed rock zone	
	regulatory	bibliography	brine chemistry	backfill	
	field	benchmark	waste packages	crushed salt	
		QA	geology	seals	
			gas flow		
			brine flow		
			mining effects		
			instrumentation		
			shaft		

2.3 Initial population of database

SITED will consolidate the results of querying several large existing databases for its initial content. The WIPP project records center has all records related to the WIPP project, including memos, data, WIPP reports, and applicable SAND reports. The Sandia Library database contains all Sandia reports (not just WIPP-related). The Office of Scientific and Technical Information (OSTI) database (<http://www.osti.gov/bridge>) has technical reports from all national laboratories. These results will be merged and imported into SITED, to form the base upon which our project-related modifications and specializations will be added. This removes a great deal of manual data entry from the project.

2.4 Electronic copies of available data

Data are stored on a variety of media, in various file formats, and of a wide range of sizes. These factors will complicate the uniform collection and archiving of both raw and processed “electronic data”. Every attempt will be made to get the data into a useful format, but time constraints will focus our efforts in this area on the obvious tests already known to have collected large amounts of data. These high-priority tests include Defense High-Level Waste (DHLW) Thermal/Structural Interaction (TSI) Rooms A, B, and H; DHLW Waste Package Performance (WPP) Rooms T and J; and Room Q.