

saple

Sandia Advanced Personnel Locator Engine

**A new way to search for Sandia people
and organizations**



November 19, 2009

Mike Procopio, Ph.D.

mjproco@sandia.gov

Cara W. Corey

cwcorey@sandia.gov

Sandia Unclassified Unlimited Release SAND #: 2009-11-11-C



Significant Initiative to Improve Search at Sandia

- **Top-down directives, support, and funding to improve the way we organize and retrieve information at Sandia**
 - Support at the CIO office level
- **Modernization of search capabilities**
 - General intranet search
 - Directory search
 - Related: Directory “Landing” page, etc.
- **Leverage modern methods**
 - Computer science, algorithms focus



In Active Deployment

- SAPLE represents an early success with search infrastructure modernization
- 10,000 search queries daily

Inside  Sandia National Laboratories

saple 

[Advanced Search](#)

Sandia Advanced Personnel Locator Engine

Use SAPLE to quickly locate Sandia personnel by any combination of first name, last name, or nickname. Spelling does not matter, as long as what you type in is a reasonably close approximation.

Search Examples

Description	Example 1	Example 2
Fuzzy name matching	mike percoppeia	j fleming
Display members in an Org, Center, etc.	5527	55=
Reverse phone number lookup (Office, Fax, etc.)	845-9653	847-????
Search by Email User ID or SNL ID	mjproco	snlid 71569
Search by job title	Managers	Lab Staff
Advanced combination searches	Bldg 758 OAAs	5??? Directors
	Division 5000 Senior Scientists	

[Show more examples and search tips](#)

maintained by  SWIFT

Website Contact
SWIFT | (505) 284-9827

For Technical Assistance
CCHD | (505) 845-CCHD

© 2009 Sandia Corporation | Operated by Lockheed Martin Corporation for the Department of Energy's National Nuclear Security Administration



Tabular View

Inside |  Sandia National Laboratories

sable  Results Format: Include Visitors / Others

[Advanced Search](#)

Searched for [org:89441](#) Results 1 – 14 of 14

Name	Email	Org	Office	Fax	Mailstop	Bldg/Rm	SNL ID	Type	Title
Adams, Tracey E.	tradaam	89441	(905) 284-0510		9019	TEMP/TEMP	500000	C	
Byers, Brian L.	bbiwers	89441	(905) 284-0542	(505) 844-9512	0947	885/117B	507386	C	
Calderon-Zablocki, Norma	ncalder	89441	(905) 284-0519		9019	912/138K	500000	C	
Castillo, Stephanie	smcasit	89441	(905) 284-0525	(505) 284-6844	0947	885/117C	507386	E	TEAM LEADR
Hadley, Lynda G.	lhadley	89441	(905) 284-0509	(925) 294-1449	9019	912/138G	500000	E	TECHNOLGST
Hokanson, Craig L.	chokans	89441	(905) 284-0537	(925) 294-2276	9019	912/150F	507386	E	TECHNOLGST
Lee, Vivien	vileeu	89441	(905) 284-0509	(925) 294-1449	9019	912/150L	500000	E	TECHNOLGST
Lieberman, Elaine S.	elieberm	89441	(905) 284-0509	(505) 844-9512	0947	885/114B	507386	E	ASA
Lorenz, Ann C.	alorenz	89441	(905) 284-0509	(505) 844-9512	0947	885/120	500000	C	
Rivers, Kay L.	kriver	89441	(905) 284-0509	(505) 844-9512	0947	885/114A	507386	E	STAFF-LAB
Scott, Janine	jscott	89441	(905) 284-0509		9019	912/138J	500000	C	
Sims, Brian P.	bsims	89441	(905) 284-0509	(505) 844-9512	0947	885/118	507386	C	
Tanaka-Lubensky, Carly S.	ctanaka	89441	(905) 284-0509	(925) 294-1449	9019	912/150E	500000	E	TECHNOLGST
Wallach, David L.	dwallac	89441	(905) 284-0509	(505) 844-9512	0947	885/119	507386	C	

sable 



Website Contact
SWIFT | (505) 284-9827
For Technical Assistance
CCHD | (505) 845-CCHD

© 2009 Sandia Corporation | Operated by
Lockheed Martin Corporation for the
Department of Energy's National Nuclear
Security Administration





Standard View

Inside  Sandia National Laboratories

sable

89441



Results Format: Standard

Include Visitors / Other:

[Advanced Search](#)

Searched for org:89441

Results 1 - 14 of 14



Tracey E. Adams

Email treadam@sandia.gov

Office (505) 844-0619

Location TEMP/TEMP (CA) | MS 9019

Fax

PO Contractor

Mobile

User ID [treadam \(SNL ID: 198009\)](#)

[Org 89441](#) | [WebCo: Web Solutions](#)

Pager

[Directory](#) | [Manager Lookup](#)



Brian L. Byers

Email bbyers@sandia.gov

Office (505) 294-0642

Location [885/117B \(TA1\)](#) | MS 0947

Fax (505) 844-9512

Staff Augmentation

Mobile

User ID [bbyers \(SNL ID: 207306\)](#)

[Graphic Arts](#)

Pager

[Directory](#) | [Manager Lookup](#)

[Org 89441](#) | [WebCo: Web Solutions](#)



Norma Calderon-Zablocki ("Norm")

Email nzablocki@sandia.gov

Office (925) 294-0539

Location [912/138K \(CA\)](#) | MS 9019

Fax

Staff Augmentation

Mobile

User ID [nzablocki \(SNL ID: 198008\)](#)

[Illustration Support](#)

Pager

[Directory](#) | [Manager Lookup](#)

[Org 89441](#) | [WebCo: Web Solutions](#)



Operational Goals

- **Foundational premise: Improve the efficiency and accuracy in personnel search**
 - **Faster searches – find the right person the first time**
 - Reduce time it takes to find someone
 - **Recall results that were not available before**
 - Find people that couldn't be found before
 - **Expose advanced search functionality more easily**
 - Natural language combination searches, like “v* building 758 oaa”
 - **More holistic and integrative search experience**
 - Integrate maps, user profile page, department homepage
 - **Improve overall user experience**



SAPLE Uses Modern Search Architecture

Traditional: Queries **Direct to DB**

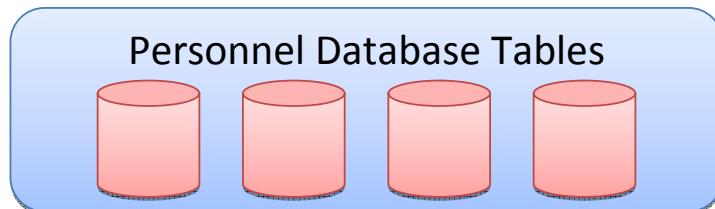


Search by [Name](#)

Search by [Nickname](#)

Search by [Organization](#)

```
select * from person WHERE name like 'pro%';
```



SAPLE: Queries Against an **Index**

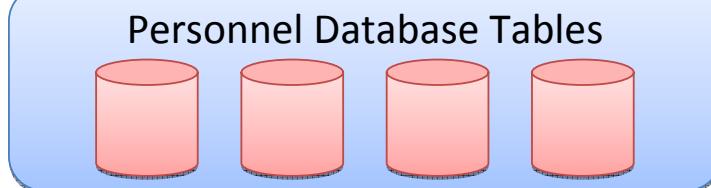
saple

Search 

[Advanced Search](#)

Artificial Intelligence and
Search Algorithms

Search Index Held in Server Memory





Three SAPLE Cornerstones

1. Inexact (approximate) string matching
2. Intelligent, compound query interpretation
3. Support of search analytics to improve results over time



Examples of **inexact string matching** are becoming increasingly common

Google: “Bernalilo country metr cort”

Web Images Maps News Video Gmail more ▾

Google™ bernalilo country metr cort

Web Results 1 - 10 of about 668 for bernalilo country metr cort.

Did you mean: [bernalillo county metro court](#)

[Bernalillo County](#) Metropolitan Court
www.metrocourt.state.nm.us/ - [Similar pages](#)

[MapMuse](#) - Locate **CORT** Furniture Rental Near You
Locate CORT Furniture Rental near you on an interactive map and directory. ... Cottonwood, AZ, Council Bluffs, IA, Country Club Hills, IL, Countryside, IL ...
find.mapmuse.com/brand/cort-furniture-rental - 324k - [Cached](#) - [Similar pages](#)

[bcmc-1.htm](#) - **Bernalillo County Metro Court**
Sep 28, 1998 ... The Bernalillo County Metro Court database contains civil case records, background checks, court dockets, hearing dates, dispositions and ...
www.technet.nm.net/menu/doc/bcmc-1.htm - 15k - [Cached](#) - [Similar pages](#)

Facebook: “Sharron Procopia”

facebook Home Profile Friends Inbox 4

sharron procpia

Did you mean: [sharon procopio](#)

Displaying all 5 results for: **sharron procpia**

	Name: Sharon 'Brender' Procopio Networks: UF Alum '02 Columbia Alum '03 Albuquerque, NM Friends: 137 mutual friends
	Name: Sharon Procopio



Better User Experience with One Query

Traditional: User must choose where to put query

SNL Directory  

This directory is to be used for official Sandia National Laboratories business only. Its use by unauthorized persons is prohibited.

Click field links for [HELP](#)

Search by [Name](#)

Search by [Nickname](#)

Search by [Organization](#)

Search by [Building](#)

Search by [Phone](#)

Search by [Fax](#)

Search by [Mail Stop](#)

Search by [Email ID](#)

Search by [Title](#)

Search by [Type](#) Employees Contractors

 [Corporate Computing Help Desk](#) , 845-2243, CCHD@sandia.gov

This application has been accessed **06271050** times.
Last Modified: May 17, 2005

SAPLE: Query is **automatically interpreted**

Inside  Sandia National Laboratories

saple  [Advanced Search](#)

Sandia Advanced Personnel Locator Engine

Use SAPLE to quickly locate Sandia personnel by any combination of first name, last name, or nickname. Spelling does not matter, as long as what you type in is a reasonably close approximation.

Search Examples

Description	Example 1	Example 2
Fuzzy name matching	mike percopiea	j fleming
Display members in an Org, Center, etc.	5527	55*
Reverse phone number lookup (Office, Fax, etc.)	845-9653	847-97??
Search by Email User ID or SNL ID	miproco	snlid 71569
Search by job title	Managers	Lab Staff
Advanced combination searches	Bldg 758 OAAs	5??? Directors
	Division 5000 Senior Scientists	

[Show more examples and search tips](#)

 maintained by **SWIFT**

Website Contact
SWIFT | (505) 284-9827

For Technical Assistance
CCHD | (505) 845-CCHD

© 2009 Sandia Corporation | Operated by Lockheed Martin Corporation for the Department of Energy's National Nuclear Security Administration



Search Analytics

- **Premise of Search Analytics**
 - Improve search quality over time by mining rich archive data logs of query history
- **Analyzing log data shows us:**
 - Usage trends
 - Number of queries daily, peak search load times, etc.
 - Where people are searching
 - Techweb home page, re-searching in SAPLE header/footer, external application hooks, SAPLE home page, etc.
 - What people are searching for
 - Last name only queries, wildcarded queries, org roster listings, compound queries, user IDs, manager lookups, etc.
 - Can take name-based queries and identify percentage that would have previously failed due to misspellings
 - How well SAPLE is performing computationally
 - Average search response time
 - How well SAPLE is performing algorithmically
 - Percentage of the time people found who they looked for on the first query



Challenges with Deploying a New Way to Search

- **People are resistant to change...**
 - **How do we improve search and other IT infrastructure offerings without disrupting flow of business activities?**
- **CIO supports this:**
 - **We must continue to evolve our offerings and progress our capabilities ... but change is uncomfortable. We must be cautious not to let vocal minority impede progress.**
 - **Tendency is do not improve capabilities for fear of vocal minority**
- **Quantifiable search improvement can help make the case**



Making the Case for Change

- **Existing code base and capabilities will not be sustainable forever**
 - Legacy code, languages, paradigms, interfaces, dependencies...
- **Assumptions and dependencies of code will change**
 - Changes in underlying personnel database structure; e.g., imminent retirement of legacy SNL Directory application



Lessons Learned

- **(If at all possible) don't rush to deployment – ensure that the tool is thoroughly tested first**
 - While a supportive user base can tolerate advances in functionality with minor glitches in the early stages, don't assume your users have endless sources of patience and good will
- **Be prepared for things to appear to be wrong – SAPLE exposed underlying issues with the data (strange nicknames, compound first names, data formatting)**
- **Edge cases can and will come up**
 - Searching for “Joe Nevada” shouldn't return “Nevada Test Site”
 - Fast, agile software development with frequent minor updates
 - Quick turnaround time to fix bugs
 - Become comfortable with incremental, minor redeployments



More Lessons Learned

- It's not easy for users to change their old habits
 - Most searches are still in the “lastname, firstname” format
- Exploring the limits of the corporate computing infrastructure
 - SAPLE is a relatively high-volume application: 10,000 queries daily
 - Each query associated with server-side algorithmic computation (e.g., string matching algorithms), which can be burdensome for a server
 - Applications have many points of failure
 - Database, Kerberos authentication, network hardware,
 - And the WebLogic production server itself with 70+ other shared applications



Still More Lessons Learned

- Increased demands on developer skill set, and it's not a simple process to transfer application support from one developer set to another:
 - Simple web-app front end that crafts SQL query from discrete textbox fields, executes query, returns results in tabular form
 - C, Perl

vs.

- String matching algorithms, regular expression matching for query interpretation, advanced data structures, multi-threading, algorithm optimization, logging and analytics, advanced user interface with rich client interactions
- Java, Java Enterprise, Servlets, XML, algorithms, database design, stored procedures, optimization
- Some advanced computer science concepts