Organizational Expertise and the Sandia National Laboratories Subject Category Guide

Jessica Shaffer-Gant and Dann Barnes

Sandia National Laboratories, Albuquerque, NM

taxonomy
BOOT CAMP >>>>

Taxonomy in Action
November 6, 2017
JW Marriott
Washington, DC

jashaff@sandia.gov
dtbarne@sandia.gov

U.S. DEPARTMENT OF ENERGY

NNSA
National Nuclear Security Administration

Who We Are and What We Do

- Sandia National Laboratories is a multi-mission laboratory with over 10,000 employees working on a wide range of projects in numerous locations with primary sites in New Mexico and California

- Sandia has four mission areas: Nuclear Weapons, Defense Systems & Assessments, Energy & Climate, and Global Security

- Sandia's fundamental science and engineering core is invested in the following research foundations: Bioscience, Computing and Information Science, Engineering Science, Geoscience, Materials Science, Nanodevices and Microsystems, Radiation Effects and High Energy Density Science

Knowledge Systems Department
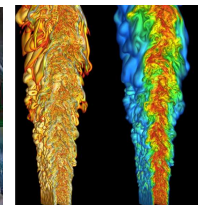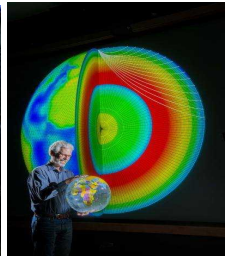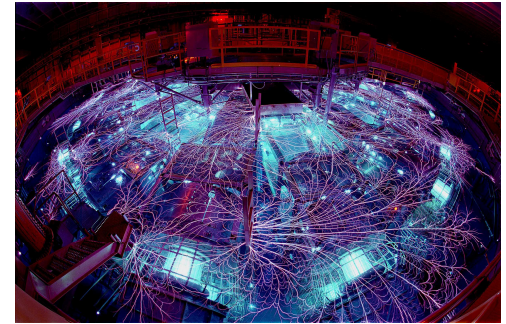
- We provide enterprise level technical solutions for information management, information analytics, and Enterprise Search

- We architect and implement innovative, high impact solutions in the field of informatics to improve the findability, usability, and protection of unstructured information for the entire Sandia workforce

ASK EXPERTISE FINDER

ASK Expertise Finder

- The Analytics for Sandia Knowledge (ASK) Expertise Finder is an application that helps users identify expertise in various disciplines across Sandia. It can be used for strategic staffing, visualization of expertise trends, and for identifying networks of collaborators

- ASK Expertise Finder indexes published documents, such as Sandia-generated reports and other information associated with individuals at Sandia. This information is produced through normal work processes; therefore, the application is self-maintaining

- The System then uses machine learning and NLP algorithms on this internally-generated data to identify people and organizations associated with a user's query term/phrase
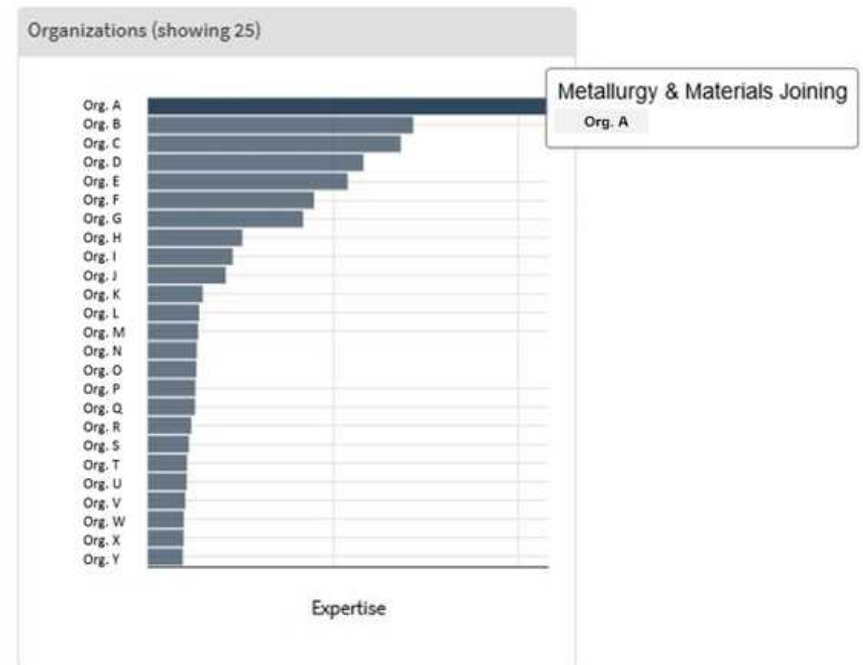
ASK Expertise Finder

ASK ORGANIZATIONAL EXPERTISE

ASK Organizational Expertise

- Expertise Finder allows us to find knowledgeable people and associated organizations based on terms/keywords

- Our management desired an extension of Expertise Finder that would organize Sandia organizations by their areas of expertise

- Using the same data in Expertise Finder and the Sandia National Laboratories Subject Category Guide (SCG), we have been able to categorize staff expertise by Sandia organization
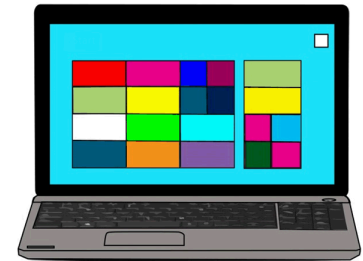
SANDIA NATIONAL LABORATORIES (SNL) SUBJECT CATEGORY GUIDE

Subject Code # ↓	Sub-category ↓	Level-2 Sub- category ↓	Subject Name ↓	Keywords and phrases ↓
25	A-J		Computers	
	A	A1 - A4	Computer Hardware and Architecture	
	B	B1 - B6	Computer Systems	
		B - 1	Desktop Computer Systems	personal computers; thin client; workstations
		B - 2	High Performance Computer Systems	concurrent computing; extreme-scale computing; high performance computing; supercomputing

SNL Subject Category Guide

- The *Sandia National Laboratories Subject Category Guide (SCG) wa*s originally based on the *Thesaurus of Engineering and Scientific Terms* (1st. ed. 1964)

- The SCG was updated and expanded in 2003 to better classify library materials in Sandia's Technical Library's online bibliographic database

- In 2013, the Materials Science areas of the SCG were expanded for use in our Knowledge Systems Department for the Competency Project; a precursor to the ASK Expertise application

- In 2016, additional subject areas were added and updated to include a second level of subcategories for Computers (primarily)

- Currently the Guide categorizes Sandia's work into 30 major subject fields broken down into 217 subcategories. Twelve of those 217 subcategories contain an additional 55 level-2 subcategories

- Each subject subcategory contains a set of pertinent terms and phrases used to match data in Sandia resources

ASK ORGANIZATIONAL EXPERTISE RESULTS USING THE SNL SUBJECT CATEGORY GUIDE
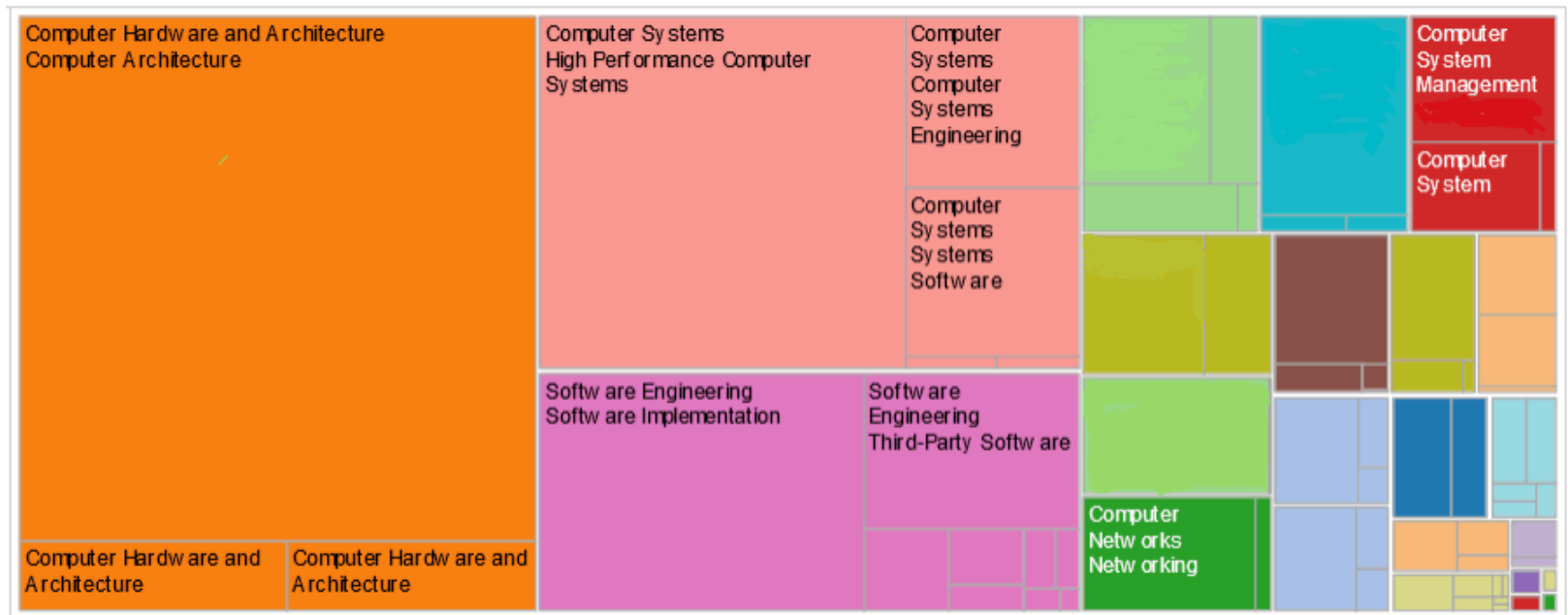
Sandia Subject Category Guide Structure (11)

Subject Code # ↓	Sub-category ↓	Subject Name ↓	Keywords and phrases ↓
11	A - M	Materials Science	
	A	Adhesives and Seals	binder; bonding; caulking; cement; epoxy
	B	Ceramics, Refractories and Glasses	alumina; oxide, silicon nitride; pzt
	↓	↓	↓
	F	Metallurgy and/or Metallography	alloy; brazing; copper; Pb; solder; steel; welding

Org Expertise Result for the "Metallurgy & Materials Joining" Department
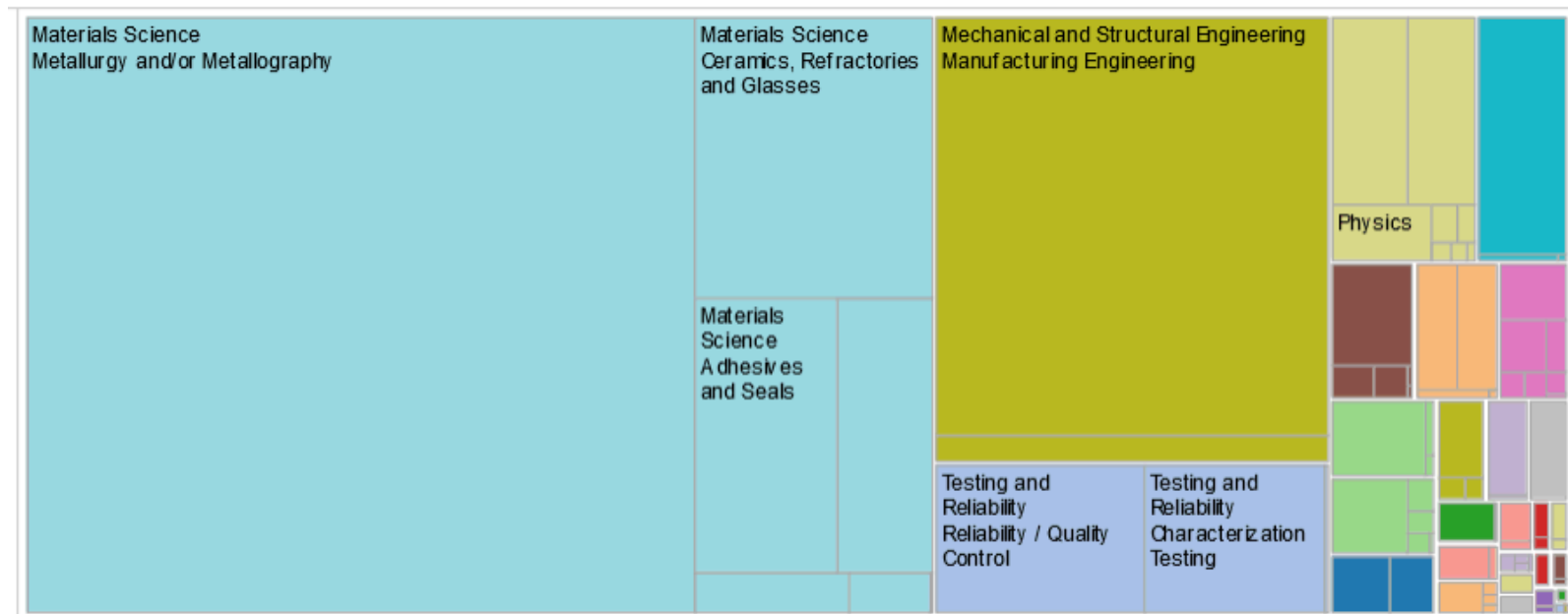
SCG Structure with Level-2 Subcategories (05)

Subject Code # ↓	Sub-category ↓	Level-2 Sub-category ↓	Subject Name ↓	Keywords and phrases ↓
05	A - I		Behavioral and Social Sciences	
	B		Documentation and Information Technology	
		B - 2	Library Science	cataloging; library services; research librarian
		B - 3	Records Management	corporate historian; NARA; recorded information
		B - 4	Information Science - General	information management; information technology

Org Expertise Visualization for the "Technical Library" Department

CHALLENGES AND FUTURE WORK TO IMPROVE SCG AND ORG EXPERTISE RESULTS

Problems and Challenges

- Problem Keywords:
  - Too general
    - performance, planning, laboratories
  - Two words of a phrase are being indexed separately or single letters with or without a hyphen are being ignored
    - C Shell, IT Security, Z machine, Control-M, Objective-C
  - Keyword has more than one meaning
    - benefits, chips, impact, lead, memory, tools
  - Acronyms are being matched as parts of words
    - MAN, RAD, SAN

- To successfully expand the Guide, assistance is needed from Subject Matter Experts (SMEs) that have the time to help provide keywords in their area of expertise

Future Work

- Analysis and implementation of an appropriate cut-off for matching scores in the Organizational Expertise visualization. If a score is insignificant (thus the block is small), then the corresponding subject area probably isn't an area of expertise for that organization

- Expand sub-categories and keywords with Subject Matter Expert assistance

- Continue to monitor results and modify or remove problem terms/phrases

- Utilize the SCG for other on-going analytics and search projects

Questions?