Military Transition Search

(as used in My Next Move for Veterans)

Author: Jeremiah Morris
jm@whpress.com

Date: November 18, 2012

Updated June 8, 2020

Overview

The focus of the My Next Move for Veterans site (https://www.mynextmove.org/vets/) is career exploration, providing information on civilian occupations and tools to discover and compare the variety of career opportunities. Using the Military Transition Search as a starting point, a veteran may find their most satisfying match in a career that's similar to, but not an exact equivalent for, their position in the military. Thus, a major goal of the search is to connect users to civilian career information as quickly and seamlessly as possible.

Using keyword search techniques, Military Occupational Classification (MOC) codes or titles matching the user's query are identified. Through the MOC crosswalk, the O*NET-SOC careers linked to the matched MOC occupations are collected and displayed.

While this paper focuses on the My Next Move for Veterans website, the same military crosswalk information is made available to developers through O*NET Web Services. The military-to-civilian crosswalk links described here are also featured within the O*NET OnLine website.

Military Crosswalk Data

While the Military Transition Search algorithm was developed by the National Center for O*NET Development, search results depend on military-to-civilian crosswalk data from several sources:

1. **DMDC** –Military Occupational Classification (MOC) crosswalk by the Defense Manpower Data Center (https://www.dmdc.osd.mil/).

2. **VOW** – An enhanced analysis (http://wdr.doleta.gov/research/keyword.cfm?fuseaction=dsp_resultDetails&pub_id=2550&mp=y) for key military occupations, which was commissioned by the U.S. Department of Labor working in collaboration with the Department of Defense, per the specifications of Section 222 of the VOW to Hire Heroes Act.

3. **RAND** – A research report on Army KSAs (https://www.rand.org/pubs/research_reports/RR1719.html) by the RAND Corporation, which identified a broad range of civilian occupations based on the knowledge, skills, and abilities (KSAs) needed by Army soldiers to perform ten of the largest Army military occupational specialties (MOSs).

4. **COOL** – Army COOL (https://www.cool.army.mil), Navy COOL (https://www.cool.navy.mil/usn/), Marine Corps COOL (https://www.cool.navy.mil/usmc/), and Air Force COOL (https://afvec.langley.af.mil/afvec/Public/COOL/) projects.

5. **CITM** – Careers in the Military (https://www.careersinthemilitary.com/), a companion website to the ASVAB Career Exploration Program.

Download links for crosswalk data from these sources may be found at the O*NET Resource Center's Crosswalks page (https://www.onetcenter.org/crosswalks.html).

The DMDC file is used as the primary source of military occupation codes and titles. The file

combines several classification systems via service codes; we load data for the Army (code A), Coast Guard (C), Air Force (F), Marine Corps (M), and Navy (N, P, and S). Loaded data includes the code, title, any listed O*NET-SOC crosswalk links, and whether the code is active or obsolete.

To facilitate matches between the DMDC file and other sources (VOW, RAND, COOL, CITM), in some cases specific codes and titles from the DMDC are grouped into broader military occupations. This happens most frequently with Air Force codes. For example, the broad occupation 2A6X2 is generated from five specific occupations:

DMDC occupations:	2A612	Aerospace Ground Equipment Helper
	2A632	Aerospace Ground Equipment Apprentice
	2A652	Aerospace Ground Equipment Journeyman
	2A672	Aerospace Ground Equipment Craftsman
	2A692	Aerospace Ground Equipment Superintendent
Broad occupation:	2A6X2	Aerospace Ground Equipment

To unify the crosswalks, we first match each military code and title from the other sources to a record in the DMDC file, or to one of the broad occupations generated based on DMDC data. The matching process primarily relies on automated mappings based on service, code, and title, with rational review to resolve the remaining unmatched entries. We retain the decisions of any rational review for use in future updates.

Once matching is complete, we have one consistent set of military codes and titles, along with links to civilian occupations from the O*NET-SOC taxonomy. We discard any military occupations not linked to civilian O*NET-SOC codes, and include military-civilian links from all sources within our master crosswalk database. To maximize the potential for career exploration, when broad occupations are involved, we apply any civilian links to both specific and broad counterparts. This ensures that appropriate matches are displayed regardless of whether the broad or specific code is requested.

The process of collecting data from these sources, matching military occupations, and building a unified crosswalk is repeated regularly to keep our web applications up to date. At this writing, our military-civilian crosswalk is generally updated quarterly.

User Experience

The user is presented with a search field where they can enter a full or partial military classification code or title, and a pull-down menu where they can optionally choose a service branch. After entering a query, the next screen displays a list of related civilian careers. The user can click a career name to see detailed information, or click on an icon beside the name to see military crosswalk information for a specific career.

When a user's query matches a single military occupation linked in the VOW analysis file described above, the user interface displays additional information from that analysis. Career results may be sorted into two tables, based on whether the civilian careers match most duties of the military occupation, or some duties of the occupation or a specialty. Beside each career is displayed the minimum military pay grade usually needed to qualify. Also displayed is an icon and text indicating the amount of preparation needed to transition from the military

occupation to the civilian career. If non-VOW data sources link to additional civilian careers, these careers are displayed separately in another listing.

If the user performs a title search with few or no direct civilian equivalents, another list of careers may be shown; these careers are matched using My Next Move's standard keyword search on the query.

"Transparent" Crosswalk

The transition search is implemented using a "transparent" crosswalk procedure. First, relevant military occupations are gathered based on the search query. Second, My Next Move careers linked to those military occupations are collected and presented to the user. The intermediate crosswalk step requires no interaction from the user, streamlining the process.

Active vs. Obsolete Military Occupations

The MOC crosswalk indicates whether each MOC occupation is active or obsolete. MOC codes are often reused for unrelated titles, leading to diverse results and potential confusion for the user. Since the target audience is in the military or recently discharged, the transition search focuses on active occupations to avoid irrelevant results for most users. Obsolete occupation codes and titles are still used when active matches are not found, to ensure that the search performs well even when older codes are used.

Each My Next Move for Veterans career in the search results features a "where in the military" icon to browse that career's military crosswalk links. Active military occupations are shown first; any obsolete occupations are displayed separately in a second section.

Code Search

A single search field is used for code and title searches. The search query is first matched against codes in the crosswalk; if no matches are found, a title search is performed.

To ensure that relevant data is returned for the most common military occupations, several passes are performed in a code search:

1. Exact match against the subset of codes linked in the VOW analysis file

2. Exact match against the subset of active MOC codes

3. Exact match against obsolete MOC codes

4. Substring match against active codes which start with the search query

5. Substring match against obsolete codes

The algorithm stops at the first pass which matches one or more military occupations, taking the optional service branch into account. If matches are found at the first pass, results are sorted based on criteria in the VOW analysis file. Otherwise, matching occupations are examined in code order, and all linked careers are added to the results. Duplicate career entries are skipped, and results are presented to the user. (Up to 20 results are initially displayed, with an option to see more.)

If an exact match is made at pass 1, 2, or 3, the matched military occupation is displayed to the user above the career results. This allows the user to confirm the correct match when results are based on a specific military occupation.

Title Search

In a title search, the user's query is matched against military titles in the crosswalk database, using a process similar to My Next Move's keyword search. (See "Title Search Details" below.) This process produces a list of matches, sorted by search relevance.

As with the code search, multiple passes are performed in a title search:

1. Exact match against the subset of titles linked in the VOW analysis file

2. Exact match against active military occupation titles

3. Exact match against obsolete titles

4. Keyword search match against all titles

Passes 1, 2, and 3 use a straightforward case-insensitive match. Pass 4 uses a process similar to My Next Move's keyword search; this process is described in "Title Search Details" below. The search stops at the first pass which matches one or more military titles.

If a match is found at pass 1, career results are sorted and displayed based on criteria in the VOW analysis file. Otherwise, matches are sorted based on relevance (pass 4) or MOC code, and all linked civilian careers are added to the results. Duplicate career entries are skipped, and results are presented to the user (up to 20 are initially displayed, with an option to see more).

If fewer than 6 O*NET-SOC occupations are identified in this way, the query is also run through My Next Move's keyword search algorithm. Additional occupations are presented in a separate table, as supplemental options for career exploration.

Title Search Details

Search queries are first normalized to strip non-alphanumeric characters and remove letter case or whitespace differences. The normalized query is then checked against also-normalized crosswalk titles, in passes 1, 2, and 3 of the title search. If the search progresses to pass 4, the normalized query is used in the steps below.

Each word of the query is matched against words in the crosswalk titles. Crosswalk entries are assigned scores if a match is found:

1. If the query word exactly matches a word in the title, the entry scores 4 points.

2. If the query word's stem matches a stemmed word in the title, the entry scores 4 points. Stems are calculated using the Paice/Husk stemming algorithm (https://web.archive.org/web/20140724170659/http://www.comp.lancs.ac.uk:80/computing/research/stemming/Links/paice.htm).

3. If a word in the title begins with the query word (a substring match), the entry scores 2 points.

Any or all of these methods may match, and contribute to the score. The cumulative score is then multiplied by a factor, based on the number of titles matched by the query word. This frequency factor increases the importance of unusual words, and decreases the importance of those that occur in a lot of titles (such as "specialist" or "supervisor").

Minimum number of matching titles	Maximum number of matching titles	Frequency factor
1	4	64
5	9	32
10	24	16
25	49	8
50	99	4
100	399	2
400	n/a	1

The query is also spell-checked, and any spelling suggestions are matched using the same process. Spelling suggestions score 2 fewer points, awarding scores of 2, 2, and zero for the three methods of matching.

After all matching methods have been applied for each query word, all matching titles are sorted in descending score order. Exact matches are re-sorted to the top, regardless of numeric score.

The relevance-score order of the crosswalk results influences the order of career matches displayed to the user. Civilian occupation links for the most relevant crosswalk title will be shown before links to a less relevant title.

O*NET OnLine

The same military crosswalk information is displayed in the O*NET OnLine website. Within OnLine's Military Crosswalk Search (https://www.onetonline.org/crosswalk/MOC/), a single search field and optional branch selection pull-down menu are presented to users, just as in My Next Move for Veterans. However, instead of the "transparent" crosswalk procedure, all matching military occupations are displayed to users. Code searches use a substring match; title searches use the query processing described above. Active military occupations are displayed first, with obsolete military occupations listed separately. Following each military match, the complete set of civilian O*NET-SOC occupations is displayed.

OnLine also displays civilian-to-military crosswalk data through an occupation's Custom Report feature. In this feature, all military occupations linked to the chosen civilian occupation are displayed. The same data sources and linkages described above are used in this area of the site.

O*NET Web Services

Through O*NET Web Services (https://services.onetcenter.org/), developers can access the Military Transition Search for use in their own websites or applications. Web Services APIs return relevant civilian matches and information about the matched military occupations

where appropriate. Data retrieved through O*NET Web Services matches the most currently available information shown in My Next Move for Veterans. For more information, see the documentation for "Find careers like your military job" (https://services.onetcenter.org/reference/veterans#military).

O*NET Web Services also provides APIs to access O*NET OnLine's military crosswalk search, returning all relevant military occupation matches in a format suitable for integrating with other applications. For more information, see the documentation for "Military crosswalk search" (https://services.onetcenter.org/reference/online#crosswalk_moc).

Performance optimizations

To provide our audience with keyword search results in a timely fashion, an efficient implementation of this algorithm is important. The O*NET Center's approach uses a relational database for storage and lookups, combined with a scripting language for input normalization and coordination of the database lookups. Improvements to this implementation include:

- **Caching of individual search components:** As described above, the military search has several components, including converting a query into a set of military codes, gathering O*NET-SOC occupations relevant to those codes, and displaying relevant information) for those occupations. We employ least-recently-used (LRU) caching to retain results for each of these parts separately. This allows us to cache and quickly return information for a single military code, for instance, across different user queries. By splitting cached data into these components, more user queries are likely to take advantage of some data already in the cache.

- **Database structure and indexing:** As with any application relying on a relational database management system (RDBMS), table and index creation is very important. We have made optimizations by following general guidelines for RDBMS optimization:
    - Creating table structures to minimize the number of JOIN operations needed
    - Designing queries to return only data necessary for later steps
    - Combining and calculating data within the database, instead of retrieving and calculating in code
    - Adding indexes to cover all queries made by the application

For optimal performance, the most important lessons we have learned are to search for performance bottlenecks at all levels of your implementation, and to carefully consider the queries and patterns you see from your particular audience.