# 5

# The Role of Technology

This chapter focuses on the presentation of the O*NET database on the web for use by individuals and organizations. It begins by describing and reviewing the O*NET user interface and then describes the use of the O*NET database by outside developers. Next the chapter discusses the structure of the database and its current presentation on the web. It then discusses the development of the "semantic web" and the possibilities for development of a semantic web structure for O*NET. The next section presents possible approaches to encourage outside development of tools and applications incorporating O*NET, and the chapter ends with conclusions and recommendations.

## THE O*NET USER INTERFACE

The field of website usability is developing rapidly and now includes sophisticated methods to gather feedback from users about their interactions with websites (Butler, 2009). Although the panel lacked the expertise and time to apply such methods to a complete review of the O*NET user interface, our brief review suggests that a more in-depth study by qualified professionals in the field of usability is needed.

The O*NET web presence includes six different web domains in a variety of styles. These web domains are only loosely connected through hyperlinks, without any navigation tools that span them. They include

- O*NET OnLine (http://online.onetcenter.org/)
- O*NET Resource Center (http://www.onetcenter.org/)

- O*NET Code Connector (http://www.onetcodeconnector.org/, also at http://www.onetcenter.org/codeconnector.html)
- O*NET Academy (http://www.onetacademy.com/)
- O*NET OnLine Knowledge Site (http://www.onetknowledgesite.com/)
- O*NET Data Collection Program (https://onet.rti.org/)

In addition, the Employment and Training Administration of the U.S. Department of Labor (DOL) maintains another website, including the following pages:

- O*NET Beyond Information-Intelligence (http://www.doleta.gov/programs/onet/)
- O*NET in Action (http://www.doleta.gov/programs/onet/oina.cfm)
- O*NET FAQs (http://www.doleta.gov/programs/onet/faqs.cfm)
- O*NET Contacts (http://www.doleta.gov/programs/onet/contact.cfm)

It appears that these sites were created by separate contractors at different stages of O*NET web development without considering how each website would fit into the larger O*NET web presence. The appearance changes from site to site, with wide variations in font sizes and column layouts. Some sites provide the viewer with little guidance about how they should be used. For example, the O*NET Academy offers a series of podcasts, recorded webinars, and online training courses about O*NET. Some of these materials appear to be aimed at job seekers, and others are targeted to employers and to government and education professionals. However, the materials are all presented together, with no sorting by the role or interest of the user.

Some of the information contained in these websites is presented on static web pages, and other information is available in the form of documents and data files that can be opened and viewed or downloaded. For example, many research and technical reports are available for viewing and downloading from the O*NET Resource Center website. Similarly, although the O*NET self-assessment tools are not available for use through an interactive website, they are available for downloading from the O*NET Resource Center website. Some of these tools are available for use in print form, and others are available in electronic form for downloading and use on local computers.

The review of the O*NET web presence that follows suggests that the multiple websites do not conform to three key principles of effective web design: (1) navigation across websites, (2) user-centered design, and (3) interactivity (Hunt, 2008a; Ta'eed, 2009). The discussion below supports

our recommendation that DOL undertake a comprehensive evaluation of all elements of the websites by usability professionals.

### Navigation Across Websites

According to Hunt (2008a) and Ta'eed (2009), providing users with a consistent set of navigation aids is one of the most important principles of website design. According to Hunt (2008a), providing navigation aids helps people know "where they are, where they can go, and gives them the means to get there easily."

Navigating through the various O*NET websites can be time-consuming and difficult. Navigation links change dramatically from one web page to another, and the user may not be sure where on the multiple sites the current page resides. For example, no links or navigation aids are provided to allow a user to easily move between the O*NET data collection page of the O*NET Resource Center website (http://www.onetcenter.org/dataCollection.html) and the O*NET data collection page maintained by RTI (https://onet.rti.org/). There is no single O*NET home page; instead, each of the O*NET web domains has its own home page.

Navigation throughout the entire O*NET web presence is also made difficult by the lack of a single search tool. To search across all the websites, the user must leave the site and go to an outside search engine, such as Google. As noted above, a great deal of information, including the O*NET database itself, is made available free for downloading by users. However, in an effort to track usage, the O*NET Academy requires users to register before gaining access to online training materials, which may discourage some users.

Some of the O*NET websites include an area labeled "O*NET Sites" displayed on the upper right and lower right of each web page, with a drop-down menu providing direct links to other websites. In general, however, our review indicates that the O*NET websites provide few aides to facilitate navigation by users. This aspect of the design of the websites stands in contrast to the research on web design reviewed by the panel, which emphasizes that ease of navigation strongly influences both the extent of use by an individual visitor to the site and the number of repeat visits.

### User-Centered Design

A second key principle of effective website design entails identifying the likely audience for the site, designing for that audience, and testing its effectiveness with that audience (Butler, 2009). Experts in website usability agree that it is essential to understand who the user is and target the site content to the user (Butler, 2009; Hunt, 2008a, 2008b; Ta'eed, 2009). If

there are different groups of users, then the web designer should design the site to meet the needs of each user group when possible, arranging the site so that users can find the information they need in a form that is relevant to them. The following sections describe O*NET OnLine and review the website in light of this principle.

### Description of O*NET OnLine

The O*NET Center makes the database directly available to the public through its website, O*NET OnLine. The website is divided into four primary categories: Find Occupations, Advanced Search, Crosswalks, and Occupation Search (see Figure 5-1). The Find Occupations area offers drop-down menus for searching the database using different criteria, such as by "in-demand industry cluster," by O*NET descriptor, or by science, technology, engineering, and mathematics (STEM) education required. The Advanced Search area has a drop-down menu linking to a Skills Search page and a Tools and Technology Search page. The Crosswalks Search area has a drop-down menu allowing the user to choose among pages for different occupational classification. These pages allow the user to enter either the specialized code from a different classification system, such as the SOC,


**FIGURE 5-1** O*NET OnLine.
SOURCE: National Center for O*NET Development. Reprinted with permission.

or keywords to search for an occupation. The Occupation Search Area includes a single text box for searching by keyword or O*NET-SOC code.

From the Advanced Search drop-down menu, the user may select the Skills Search page. This page allows the user to select any number of Basic Skills, Complex Problem-Solving Skills, Resource Management Skills, and other skill descriptors and search for occupations that most closely match that combination of skills in the O*NET data. Alternatively, the user may select Tools and Technology. This page contains a text box for entering specific tools or technologies in order to search for occupations using these tools.

**Usability of O*NET OnLine**

The O*NET OnLine home page states that it offers "detailed descriptions of the world of work for use by job seekers, workforce development and HR professionals, students, researchers, and more!" (see Figure 5-1). However, the home page contains drop-down menus with many undefined acronyms and terms with meanings that cannot be obvious to any but very experienced users.

One example of the difficulty users may encounter is in the Find Occupations section of the home page, which invites the visitor to "Browse groups of similar occupations to explore careers." The user is offered a drop-down menu to search by Career Cluster, Green Economy Sector, In-demand Industry Cluster, Job Family, Job Zone, O*NET Descriptor, or STEM discipline. To respond to this menu, a novice user would first be required to look up the meanings of these terms. Similarly, the Crosswalks Search area offers a drop-down menu of the following options: CIP, DOT, MOC, OOH, RAPIDS, and SOC. Nowhere on the home page is there any indication of the meaning of these acronyms. The lack of clarity in these two areas of the home page, which provide key entry points to the database, may reduce interest in and use of the website (Hunt, 2008b; Tufte, 1990).

If the user persists and opens one of the drop-down choices in the Crosswalks Search area, he or she will find the definition of the particular acronym chosen. These definitions, such as Classification of Instructional Programs for CIP and Military Occupational Classification for MOC, are unlikely to be meaningful to the average lay person (although they may hold meaning for specialists in occupational and educational classification systems).

Continuing in the Crosswalks Search area, the user finds a page for each occupational classification system with a text box for searching either by occupational title or by using that classification system's code. If the user enters a title in one of these search boxes, the website returns a page with a long list of links to occupations. The page does not provide definitions of

the occupations. Selecting an occupation from this list leads to the summary report—one web page containing approximately four printed pages of text in a small font. Research evidence in the field of usability suggests that users often do not read web pages presented in this way (Hunt, 2008a, 2008b; Ta'eed, 2009). A workforce development specialist who spoke to the panel described the level of detail in the summary report as "overwhelming for the lay person" (Graybill, 2009).

A user who chooses the detailed view for an occupation receives approximately 10 pages of text listing all of the 200-plus descriptors in the content model, with bar graphs showing the importance of each descriptor (Skills, Abilities, etc.). Comparing these characteristics of the occupation with those of another occupation requires the user to first print the pages associated with one occupation, then back out and select the other occupation and print out those pages. As an alternative, the user could save each subsection of each page to a separate worksheet program (such as Microsoft Excel) or to a generic comma-separated-values file. The website's lack of support for side-by-side comparisons of occupations hinders workforce development specialists from analyzing skill gaps of displaced workers or designing customized retraining programs to fill those gaps (Graybill, 2009).

The panel's brief review, together with the lack of any data based on user testing or tracking user interactions with the website suggests that development of O*NET was not based on the principle of user-centered design. The design does not reflect a process of identifying the likely audience for the site, designing for that audience, and testing the effectiveness with that audience.

### Interactivity

The third principle of modern web design is to make information service interactive and responsive to user input (Hunt, 2008a, 2008b; Ta'eed, 2009). O*NET OnLine limits the user's ability to assemble and compare similar data across multiple occupations. For example, a user viewing the detailed information about an occupation, including the importance levels for various descriptors of Skills and Abilities, cannot easily view other occupations for which these descriptors are equally important. Although it is possible to select some occupational characteristics (Skills, Abilities, Work Values, and others) and to search for occupations requiring these characteristics, it is not possible to search on the basis of other criteria, including required level of education and median wages. Nor can a user select classes of occupational characteristics and then narrow selection according to different characteristics. For example, the user cannot identify the set of oc-

cupations for which knowledge of psychology is important and then select the subset that requires a master's degree.

The next generation of Internet tools, often referred to as Web 2.0, allows website users to do more than simply retrieve information. These tools divide data storage from its presentation and provide users with great flexibility in how they use data (Schroeder et al., 2009). Similarly, the use of scripting languages, backend data storage, and flexible display design can make the user's interaction with a website more productive. Application of these tools could encourage use of O*NET OnLine.

## USES OF THE O*NET DATABASE

In addition to making the database available to the public through O*NET OnLine, the O*NET Center makes it available for download and use by outside organizations and developers. Many individuals and organizations download the database and incorporate it into tools and applications that are used in career development, workforce development, human resource management, and research (National Center for O*NET Development, 2009). From January 2002 to December 2008, the database was downloaded over 70,000 times (Lewis and Rivkin, 2009a). Although the database is free and no license is required, the O*NET Center requests developers to comply with a formal user agreement (see Box 5-1).

To cite just one of many examples, the University of Oregon downloads the O*NET database to create occupational profiles that are linked to other data on education and training programs and financial aid in a comprehensive career information system (CIS). This system is used in more than 7,000 schools and state agencies, and it serves as the basis for state-based career information delivery systems in 15 states (see Chapter 6). The University of Oregon CIS is one of 14 large online career guidance systems that incorporate the O*NET database; these systems provided O*NET data to approximately 37 million individuals in February 2009, with the typical user accessing the database throughout academic year 2008-2009 (Janis, 2009).

By comparison, O*NET OnLine received 766,000 visitors per month in 2008, or approximately 9.2 million visitors over the course of 2008 (Lewis and Rivkin, 2009a). These 9 million visitors carried out 10.5 million searches and viewed more than 13 million occupational reports, translating to just over one search per visitor and one report per visitor (Lewis and Rivkin, 2009b). Most activity at the O*NET OnLine website was concentrated in the keyword search area.

On the basis of the limited evidence available, it appears that the use of O*NET OnLine is much smaller than the use of O*NET by visitors to external websites incorporating O*NET data. While it is important to

**BOX 5-1**
**User Agreement and Certification**

Developers of products, software, or system applications using O*NET are subject to the terms of a formal O*NET Database Products User Agreement (http://www.onetcenter.org/agree/database). The agreement requests the individual to complete a registration form with contact information that is sent to the O*NET Center. The registration form asks the individual about his or her willingness to share information about their use of O*NET products; those who agree to this condition provide valuable information on usage to the O*NET Center and to DOL. However, some of those who download the database elect not to complete the voluntary registration form, and some of those who complete it elect not to be contacted to share information on their use of O*NET.

The terms of the user agreement include acknowledging that O*NET® is a trademark of the U.S. Department of Labor, Employment and Training Administration, and displaying the trademark symbol. The user must also display the version number of O*NET included in the product, such as O*NET 14.0, and must display the "O*NET In It" mark in and on the product. To support this element of the user agreement, the O*NET Center makes the "O*NET In It" graphic available for download and reproduction (http://www.onetcenter.org/graphics.html). In addition, the O*NET Center provides O*NET OnLine graphics for use by website developers that link to O*NET OnLine from within their own websites and graphics.

maintain a user interface providing direct access to the database, this function should be seen as clearly secondary to the task of maintaining the core O*NET data. Any web development efforts should focus on the provision of services that are not effectively offered elsewhere. The finding that the database is used most widely in applications created by outside developers leads to our discussion in the following sections of the role of technology both to increase availability and use of O*NET and to assist in gathering background occupational data.

## STRUCTURE AND AVAILABILITY OF THE DATABASE

To encourage outside developers to make use of the most current O*NET database, referred to as the development database, the O*NET Center makes it available for download approximately 3 months before it is incorporated into O*NET OnLine. For example, the O*NET Center made O*NET 14.0, the current development database, available for download

in June 2009.[1] At that time, O*NET OnLine continued to incorporate the older version of the database, O*NET 13.0, which the O*NET Center refers to as the production database. The most current database, O*NET 14.0, was incorporated into O*NET OnLine in fall 2009.

### Database Content

The database includes, for each occupation, survey information related to most domains in the content model, such as descriptors of Abilities, Skills, and Work Values.[2] Although information on workforce characteristics, one of the six major categories of information in the content model (see Figure 1-1) is included in O*NET OnLine, it is not included in the database made available for download. In addition to the database comprised of core files related to the content model, the O*NET Center makes other information, including Emerging Tasks, Detailed Work Activities, Tools and Technology, and crosswalk files available separately for download. The O*NET Center refers to these other data sets as "supplemental files." It maintains archives of all survey data collected and earlier published versions of the database and the supplemental files.

### Availability of the Database

Users can directly download the core O*NET data from the O*NET Resource Center website in the format of a compressed file containing 22 separate files stored as text. The supplemental files, such as the Detailed Work Activities, are also made available as separate, compressed text files. Each data table is contained in one of the 22 separate files, and assembling the full database is cumbersome. The user must download each file separately, uncompress it, and read it into the user's application. The user must specify the fields in a table to be used to uniquely identify each observation and must also specify the relationships between rows in different tables. Users may not be aware that an "O*NET Data Dictionary" is made available along with the database, or that this dictionary includes a section describing the relationships between the tables and the elements of the tables that operate as identifiers. The complex and time-consuming process for downloading the database and the supplemental files could easily result in errors.

[1]While making the most current database available for downloading, the O*NET Center also continues to make the previous database available for those who prefer that version.

[2]In addition to the mean value of survey responses, the O*NET tables include a few additional items of summary statistical information about the data collected to populate the content model. Some additional detail on the survey instrument responses is provided in a table called Occupation Level Metadata.

Downloading the database and supplemental files is easier for users whose computers employ the Windows operating system. The National Crosswalk Service Center supported by the Department of Labor provides the O*NET data preloaded for three proprietary database programs: Microsoft Access, Visual Fox Pro, and SAS, all Windows applications with proprietary data formats.

Downloading the database would be easier for all users if it were made available in structured query language (SQL) format. SQL is a nonproprietary computer language that can be used to query and modify data and manage databases. It is designed specifically to facilitate the retrieval, insertion, updating, and deletion of data in relational databases, such as O*NET. It is widely used and has been standardized by the American National Standards Institute (ANSI).

Making O*NET available in ANSI SQL format would facilitate downloading by developers who use the three currently supported database programs. It would also facilitate the downloading and use of O*NET by developers using other proprietary and nonproprietary database programs. Making the database available in ANSI SQL, in addition to the currently available versions, would lower developers' costs of creating local copies and reduce the probability of error for developers who do not use the three currently supported database programs.

## THE SEMANTIC WEB

Much of the usefulness of O*NET is derived from linking it to other data sets, including data on wages and employment levels, to create new applications and tools (see Chapter 6). New technological tools could facilitate links between O*NET and other data sets on the Internet, expanding the potential for developers to create useful applications. This section introduces an emerging technology for linking data sets across the Internet: the semantic web.

The early development of the Internet was motivated by the need to efficiently transmit large quantities of data between locations (Berners-Lee, Hendler, and Lassila, 2001). The early World Wide Web incorporated a presentation protocol that allowed transmission of human-readable information between computers. Over time, the increasing use of data in web applications led to the development of tools, most notably XML, for transmitting data in a way that preserved its structure and allowed greater ease in communicating it between computers as long as the sender and recipient agreed on the meaning of the content of the data. These tools are now being used to exchange occupational information.

### The HR-XML Consortium

Technical experts in the human resources field have created standards for communicating data between computers. The HR-XML Consortium (http://www.hr-xml.org) is made up of over 70 members, including Monster, SAP, Microsoft, Oracle, Kelly Services, ADP, EDS, and the Society for Human Resource Management. The consortium publishes and develops XML schemas for transaction processing in human resource departments and firms. It has used O*NET data to identify common elements of human resource management systems and develop specifications that allow organizations to capture and use occupational information for human resource functions, such as building competency models, conducting job analyses, and developing performance appraisal systems (U.S. Department of Labor, 2008).

Standards developed by the HR-XML Consortium enable automation of human resource–related data exchanges. For example, through HR-XML, a standardized set of information can be exchanged between a staffing supplier (such as a staffing agency) and its customer (such as a manufacturing firm). This information might include order processing, key competencies of employees, employee assignments and hourly rates, hours worked, and payroll expenses. The HR-XML communication protocol may be seen as a step in the direction of an ontology for the field of human resources (Bizer et al., 2005), and an ontology for HR-XML may not be that far off (Bohring and Auer, 2005).

Although XML provides a reliable way to transmit data, it does not provide significant "semantic content," or a definition of the data independent of the processes being used to transmit them that would allow computers to read and interpret the data. A database such as O*NET can be given meaning by defining an *ontology* for it. An ontology is a formal representation of meanings and relationships in a given context, in this case, data stored on the web. The main standard-setting body for protocols used on the World Wide Web has created a system of semantic web standards (World Wide Web Consortium, 2009), designed to provide information about the meaning of a data element in a way that can be interpreted unambiguously by a computer program. This system, the Resource Description Framework (RDF), assigns a specific context to web objects (Powers, 2003). The key standards for expressing ontological information are the RDF Schema (RDFS) standard and the Web Ontology Language (OWL) (Allemang and Hendler, 2008). It is important to note that, although the use of these current standards is growing, new semantic web tools may emerge in the future.

### Standards for the Semantic Web

In the RDF system, the label used to assign an unambiguous context to an object or data set on the web is referred to as a Uniform Resource Identifier (URI). A URI is simply a web address that is a global identifier for the ontology of the data that are being made available on the web. The ontology allows a computer program to draw inferences about meanings and relationships of objects in the data set.

Data that can be retrieved through a URI are often referred to as an RDF store. As the owners of various data sets establish RDF stores, it becomes possible for data users, including researchers and applications developers, to write a data query that draws data from distinct, but semantically linked, data sets (Sadler, 2009). Such a query can select, filter, and merge data from the various data sets. Semantically linking data sets offers several benefits to the owners of these data sets and to data users generally. It allows Internet-wide data mining and analysis, provides the opportunity for interoperability between widely dispersed data sources, and increases the likelihood that unexpected innovations using linked data can develop (Sadler, 2009). Recognizing these benefits, the owners of many large public data sets on the web have assigned ontologies to their data sets, and these data sets are now semantically linked (see Figure 5-2).

### Semantic Web Tools in Federal Agencies

The federal government has recently begun to deploy semantic web techniques. For example, Jet Propulsion Laboratories, under contract with the National Aeronautics and Space Administration (NASA), has developed and is continuing to expand ontological schema for all data derived from NASA planetary exploration programs. Jet Propulsion Laboratories is also building an ontology for environmental science information, designed to allow researchers to remotely query any of the numerous data stores distributed around the country, using different computer platforms and different storage technologies. The National Cancer Institute has released a public ontology that allows computers to more easily assist in exploring information on genetics, biology, and chemistry developed through institute-sponsored research and other research in order to advance the institute's mission.

The Office of Management Budget has developed an ontology for its Federal Enterprise Architecture initiative, a government-wide effort to maximize the contribution of information technology systems and resources in support of each agency's performance goals.

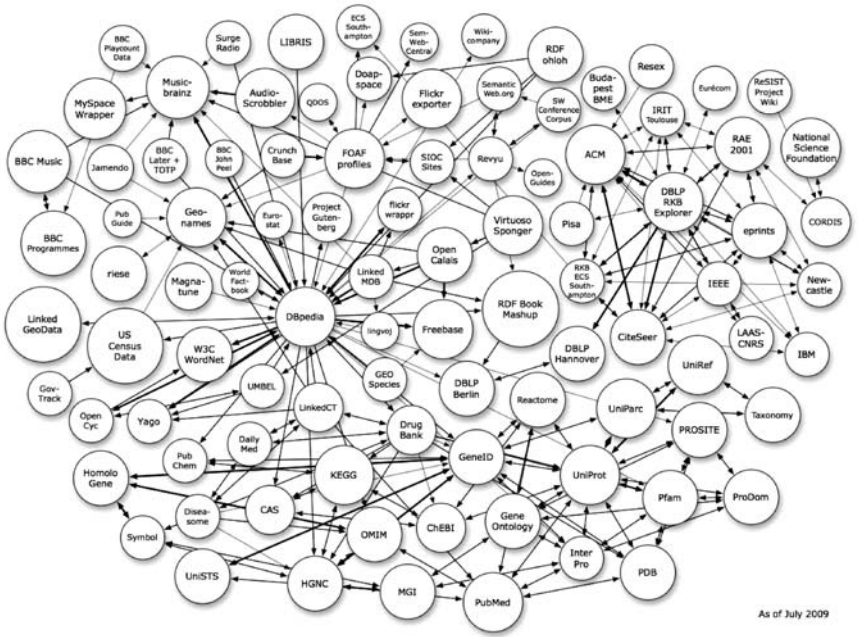


**FIGURE 5-2** Large public data sets that are semantically linked on the web. SOURCE: Cyganiak and Jentzch (2009). Reprinted with permission.

## SEMANTIC WEB POSSIBILITIES FOR O*NET

As the creators of the data, DOL and the O*NET Center are uniquely positioned to define authoritatively the ontology of O*NET data. If DOL created an ontology for O*NET and then placed the structured database on the web, web users could query the O*NET data, along with any other related, semantically enabled data, and use them for display, for research, or for local storage to serve as an input to a local application. In this scenario, any user with a web browser would be able to query the data, using SQL or another computer language or application.[3] For example, a job-seeker using her laptop at a coffee shop could use her own copy of Microsoft Excel, with its built-in database query tool, to query the O*NET data. A state employment office might develop a small application that would allow the laptop user at the coffee shop to run a program that queries O*NET

[3]As we discuss below, it is not necessary that the data image being queried is on a server operated by DOL or the O*NET Center. A public data image on a DOL or Center server may or may not be one of the data stores responding to semantic queries.

and combines the result with online information about local job openings and places markers with phone numbers on an online map, together with local bus routes and current bus locations. Such an application would draw on semantic data from at least four different sources, combining them to deliver a useful service.

In this scenario, the O*NET Center would continue its current practice of publishing database images for download by users who need to keep a local copy of the data. The key difference is that, with a clear public ontology, the O*NET database images would not only be used privately as an input to a local application, but would also be accessible from other public sites. This service redundancy for Internet data is often referred to as "mirroring."

The structure of semantic relationships in the O*NET database is relatively simple, because the data are already organized using a classification system based on the SOC and the content model. Creating an ontology to describe these existing relationships between data elements in O*NET would not require a large amount of time or money.

According to Bizer et al. (2005), semantic linkage between O*NET and other human resource data sets might result in a substantial expansion of the use of O*NET in human resource management information systems and in human resource management processes more generally. If the O*NET data and its ontology were published to the web, DOL might set up a process to encourage development of innovative applications of O*NET.

### Possibilities for Semantic Web Linkages

If it developed an ontology for O*NET, DOL might want to extend it to include related data sets. As noted above, many third parties currently download the O*NET database, link it with other data, and then publish the result in a local application or service. These efforts might accelerate through semantic web linkages between O*NET and other occupational and education data sets.

The National Crosswalk Center has already developed crosswalks between O*NET and other occupational and education data sets, including the SOC and the Classification of Instructional Programs. Adding these crosswalks to a new O*NET ontology would allow computers to understand the relationships between O*NET and these crosswalks. Including the SOC in the ontology would facilitate automated discovery (data mining) and linking of O*NET and SOC data and allow further linkages to be made to any data with a computable relationship to SOC.

### Wiki Technology

Later in this report, we identify weaknesses in communication between the O*NET Center and O*NET users (see Chapters 6 and 7). Deployment of online collaboration tools, or wikis, could help DOL and the O*NET Center enhance communication with users and with outside experts in occupational analysis, data collection, and information technology.

The most famous and widely used wiki tool is Wikipedia, the collaboratively created online encyclopedia. Wikipedia Foundation provides the computing infrastructure, the server, wiki software, general rules for entries, and style guidelines. Content is generated by anyone who has access to an Internet browser. Users can edit existing content pages or create new pages on topics not yet covered. The Wikipedia wiki software provides the online editing environment, tracks the changes made to pages, and allows contributors to engage in an online discussion about the content of pages. Page and text formatting is accomplished by simple specialized mark-up tags.

Wiki software tools have been adopted by corporations, universities, and other organizations to encourage sharing of best practices and exchange of knowledge. Federal agencies have also begun using these tools (Schroeder et al., 2009). For example, the U.S. General Services Administration (GSA) operates a wiki environment to encourage communication across governmental entities (see http://colab.cim3.net/cgi-bin/wiki.pl?WikiHomePage). The GSA site helps people across federal agencies who are engaged in a particular subject or project to share experiences and learn from each other through a central online point of contact, rather than attempting communication through a series of e-mail conversations.

DOL and the O*NET Center might use wikis and other online collaboration tools to establish an ongoing dialog with local workforce development agencies, career information delivery systems, human resource management associations, and other O*NET users. Such a dialog could potentially enhance service to users and quality of the database. DOL could also use wiki technology to generate a valuable two-way flow of information about technical issues between outside researchers and the O*NET Center that could inform improvements to the database and data collection methods.

Wiki technology might also be used to support the O*NET Center's collection of data related to newer domains of descriptors, such as Tools and Technology and Emerging Tasks as well as data related to new and emerging occupations. These data are not collected through national surveys of job incumbents, but are gathered primarily through Internet searches and consultations with professional associations and customer feedback, sometimes supplemented by expert review (e.g., Dierdorff, Drewes, and Norton, 2006).

Using wiki technology to establish open lines of communication with trade and professional associations and outside experts would allow the O*NET Center to obtain a broad range of information related to these domains and new occupations at low cost. However, it is important to note that wiki systems encourage all users to freely contribute information. Users, who are self-selected, may enter information and edit others' entries, making corrections or offering additional information. The resulting information would be of uneven quality and would not be nationally representative. The O*NET Center would need to conduct further review of the information and compare it with other data sources of known quality before deciding whether to include it in O*NET. The semantic web technologies described above could also assist in data collection, making it easier for professional associations and others with knowledge of occupations to provide information to the O*NET Center. However, this information, too, would be of mixed quality and would not be nationally representative.

## LEVERAGING DEVELOPMENT EFFORTS

Because most use of O*NET data takes place through applications developed outside DOL and the O*NET Center, DOL may want to consider several possible approaches to encouraging and expanding development of these outside tools and applications. These approaches might be taken separately or in combination.

One possibility would be to support the creation of an open-source development community. Communities of developers using open-source code, which is freely available on the Internet, have created major software products including web servers (Apache) browsers (Firefox), word processors (Open Office), and the Linux operating system. Major private companies, including IBM, Hewlett-Packard, Apple, Sun Microsystems, and RedHat Software also participate to some extent in the open source model of software development. DOL might set up a process to encourage and possibly reward the development of innovative, nonproprietary applications of O*NET data that would be made available to all interested parties at no charge.

Another possibility would be to offer awards or prizes to the best new applications using O*NET. The technical advisory board recommended in this report might participate in selecting the best new O*NET applications, which might include standalone applications, new related semantic web objects, web applications, mashups with O*NET mixed in, and code samples that users could copy and drop into their own applications.

The Environmental Protection Agency proposed another possible approach to supporting outside developers in its roadmap for publishing environmental data. The roadmap envisions the creation of widgets—reusable web code—that could be used by service providers to build new web pages

incorporating environmental data (U.S. Environmental Protection Agency Web Workgroup, 2008). Such an approach would provide DOL with a mechanism for measuring O*NET use by these tools.

Another possibility would be to hold a periodic conference of O*NET developers, users, and researchers for sharing new studies, uses, and applications of O*NET. A conference setting would also provide an opportunity for user feedback to DOL on how O*NET could be enhanced.

## CONCLUSIONS AND RECOMMENDATIONS

The O*NET database is incorporated into many web applications that are easy to search and are widely used. However, the formats in which the database is currently made available for download—either as a series of text files or in three proprietary programs used with the Windows operating system—limit even wider incorporation of O*NET data within these web applications.

**Recommendation: The Department of Labor should, with advice and guidance from the technical advisory board recommended in Chapter 2 and the user advisory board recommended in Chapter 6, explore methods for distributing O*NET data in platform independent, nonproprietary formats that facilitate its acquisition and use in the widest variety of database applications. In particular, DOL should make the O*NET database available in the American National Standards Institute (ANSI) structured query language (SQL) in addition to the current formats. In addition, DOL should make efforts to encourage the development of O*NET applications by states, private firms, and educational institutions. The active development of O*NET applications by others would allow DOL to focus O*NET resources on the core functions of collecting, maintaining, and publishing high-quality data.**

A review of web design literature and a preliminary evaluation of the O*NET websites suggests that the sites do not conform to three key principles of web design: ease of navigation, targeting of content to users, and interactive elements that can flexibly respond to user interests and styles of use.

**Recommendation: The Department of Labor should, with advice and guidance from the technical advisory board recommended in Chapter 2 and the user advisory board recommended in Chapter 6, consider conducting a usability study to obtain user feedback on the ease of use of the O*NET websites. The study should emphasize the basic website functions of providing access to information in the O*NET database**

**and to information about O*NET. Toward this end, the study should identify likely users of the site and lead to improvements that will provide these users with easy ways of navigating to information that is relevant to their particular needs. The goal should be to present information on the websites in a way that is engaging and encourages exploration.**

The use of Web 2.0 or wiki-type processes in which those who retrieve information from a website are also able to add to and update some items interactively online could potentially support wider dissemination of the O*NET database and enhance collection of information on occupations. However, these processes would not yield nationally representative data.

**Recommendation: The Department of Labor should, with advice and guidance from the technical advisory board recommended in Chapter 2 and the user advisory board recommended in Chapter 6, commission research exploring the potential for using wiki processes to enhance communication with and among O*NET users and to obtain occupational information. However, data gathered using these processes should not be a replacement for data collected systematically from carefully specified samples.**

Semantic web techniques have the potential to enhance uses of O*NET data and significantly expand the opportunities to link O*NET data to other data sets available on the web.

**Recommendation: The Department of Labor should, with advice and guidance from the technical advisory board recommended in Chapter 2 and the user advisory board recommended in Chapter 6, explore semantic web techniques. Specifically, DOL should consider creating an occupational classification ontology that would encompass O*NET and the Standard Occupational Classification system. The usefulness of this new ontology would be enhanced by continuing efforts in the future to add semantic links to closely related data sets, such as data sets on education and training programs and workplace competencies.**

## REFERENCES

Allemang, D., and Hendler, J. (2008). *Semantic web for the working ontologist: Effective modeling in RDFS and OWL*. Amsterdam: Morgan Kaufmann.

Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The semantic web. *Scientific American Magazine*. May.

Bizer, C., Heese, R., Mochol, M., Oldakowski, R., Tolksdorf, R., and Eckstein, R. (2005). The impact of semantic web technologies on job recruitment processes. In O.K. Ferstl, E.J. Sinz, S. Eckert, and T. Isselhorst (Eds.), *Wirtschaftsinformatik 2005: eEconomy, eGovernment, eSociety* (pp. 1367-1381). Heidelburg: Physica-Verlag HD.

Bohring, H., and Auer, S. (2005). Mapping XML to OWL ontologies. Abstract available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.59.8897 [accessed July 2009].

Butler, S. (2009). *Improving the usability of O*NET*. Presentation to the Panel to Review the Occupational Information Network (O*NET). Available: http://www7.nationalacademies.org/cfe/O_NET_Scott_Butler_Presentation.pdf [accessed May 2009].

Cyganiak, R., and Jentzch, A. (2009). *The linking open data dataset cloud*. Available: http://richard.cyganiak.de/2007/10/lod/ [accessed August 2009].

Davis, S. (2007). *GIS for web developers: Adding "where" to your web applications*. Lewisville, TX: Pragmatic Bookshelf.

Dierdorff, E.C., Drewes, D.W., and Norton, J.J. (2006). *O*NET tools and technology: A synopsis of data development procedures*. Raleigh, NC: National Center for O*NET Development. Available: http://www.onetcenter.org/reports/T2Development.html [accessed July 2009].

Graybill, B. (2009). *Occupational skills and the workforce*. Presentation to the Panel to Review the Occupational Information Network (O*NET). Available: http://www7.nationalacademies.org/cfe/Graybill%20Power%20point.pdf [accessed July 2009].

Hunt, B. (2008a). *Save the pixel: The art of simple web design*. Available: http://savethepixel.org/save_the_pixel_ch1.pdf [accessed July 2009].

Hunt, B. (2008b). *Web design from scratch: Basics*. Available: http://webdesignfromscratch.com/basics.php [accessed May 2009].

Janis, L. (2009). *O*NET uses in career guidance*. Presentation to the Panel to Review the Occupational Information Network (O*NET). Available: http://www7.nationalacademies.org/cfe/Janis%20Power%20point.pdf [accessed July 2009].

Lewis, P.M., and Rivkin, D.R. (2009a). *O*NET program briefing, February 3*. Presentation to the Panel to Review the Occupational Information Network (O*NET). Available: http://www7.nationalacademies.org/cfe/Rivkin%20and%20Lewis%20ONET%20Center%20presentation.pdf [accessed July 2009].

Lewis, P.M., and Rivkin, D.R. (2009b). *General user feedback on O*NET*. Presentation to the Panel to Review the Occupational Information Network (O*NET), March 26. Available: http://www7.nationalacademies.org/cfe/Rivkin%20and%20Lewis%20Power%20point.pdf [accessed March 2009)]

National Center for O*NET Development. (2009). *O*NET products at work*. Raleigh, NC: Author.

Powers, S. (2003). *Practical RDF*. Sebastopol, CA: O'Reilly.

Sadler, E. (2009). *Small pieces loosely joined: An argument for publishing O*NET as linked data*. Presentation to the Panel to Review the Occupational Information Network (O*NET). Available: http://www7.nationalacademies.org/cfe/O_NET_Elizabeth_Sadler_Presentation.pdf [accessed July 2009].

Schroeder, D., Joubert, D., Meyer, P., Steinhauser, L., and Walker, A. (2009). *Report and recommendations on web 2.0 and social software federal knowledge management working group of the federal knowledge management initiative*. Available: http://wiki.nasa.gov/cm/wiki/?id=7733 [accessed December 2009].

Ta'eed, C. (2009). *Essential principles for good web design*. Available: http://psd.tutsplus.com/designing-tutorials/9-essential-principles-for-good-web-design/ [accessed May 2009].

Tufte, E. (1990). *Envisioning information*. Cheshire, CT: Graphics Press.

U.S. Department of Labor, Employment and Training Administration. (2008). *O*NET data collection program, Office of Management and Budget clearance package supporting statement, Volumes I and II*. Raleigh, NC: Author. Available: http://www.onetcenter.org/dl_files/omb2008/Supporting_Statement2.pdf [accessed June 2009].

U.S. Environmental Protection Agency Web Workgroup. (2008). *Environmental Protection Agency Web 2.0 whitepaper*. Washington, DC: Author.

World Wide Web Consortium. (2009). *W3C semantic web activity*. Available: http://www.w3.org/2001/sw/ [accessed May 2009].