

A multilingual knowledge management system: A case study of FAO and WAICENT

Daniel E. O’Leary

University of Southern California, Los Angeles, CA 90089-0441, United States

Available online 27 July 2007

Abstract

This paper provides a case study of a multilingual knowledge management system for a large organization. In so doing we elicit what it means for a system to be “multilingual” and how that changes some previous research on knowledge management. Some researchers have viewed multilingual as meaning a multilingual user interface. However, that is only a small part of the story. In this case we find multilingual also refers to a broad range of “multilingual,” including multilingual knowledge resources, multilingual feedback from users, multilingual search, multilingual ontologies and other concerns.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Multilingual system life cycle; Knowledge management; Multilingual Ontologies; Controlled Vocabularies; Evolution; Multilingual Search

1. Introduction

In the early and middle 1990’s the Internet’s first web pages were virtually all in English, as were the search engines [23]. However, that has changed as web has users from all over the world now access the Internet, and web sites are available in virtually every different language. In some cases specific organizations must, or simply find it important and cost beneficial to accommodate those multilingual users.

In order for different groups to use the Internet, and its far reaching access to use available knowledge resources, the knowledge resources need to be available in multiple languages. In particular, large, multinational corporations, and many governments and governmental organizations require the use of multilingual systems for their knowledge resources. For example, in Belgium four languages are in use: Dutch, French, German and English (<http://www.belgium.be/eportal/index.jsp>). Fur-

ther, important developments whether medical or governmental occur or are consumed in many languages [26]. As a result, knowledge resources must be accommodated and promoted in multiple languages.

Unfortunately, there are few models or case studies in the literature available to illustrate the use of multilingual knowledge management systems. Further, it is not clear what the emerging architectures are for the development of such multilingual systems. As a result, it is also not clear what are some of the key or emerging issues associated with developing, implementing or maintaining multilingual systems.

1.1. Purpose

Accordingly, the purpose of this paper is to develop a case study of a multilingual knowledge management system. In so doing we are able to analyze some of the approaches used in the development of a multilingual capability. Further, such a case study allows us insight into what are some of the primary multilingual capabilities

E-mail address: oleary@usc.edu.

required or used in such systems and how those capabilities are implemented. In addition, such a case study can provide us with insights into some of the problems and decisions associated with such systems.

An analysis of the literature reveals few such case studies of “real world” organizations. To-date, much of the literature is aimed at addressing specific problems (library resource overload), generally using a particular technology (e.g., portals) and does not provide a portfolio of multilingual solutions that are part of such real world organizations. In addition, this case study provides a benchmark, from which we can gauge the growth of multilingual capabilities. Further, by analyzing what multilingual activities are employed in a real world case study we can better understand what it means to provide a “multilingual” system. Finally, comparing multilingual knowledge management capabilities to other knowledge management systems lets us note that previous investigations into knowledge management have been “underspecified.”

1.2. *This paper*

This paper proceeds in the following manner. Section 2 provides a review of some of the multilingual knowledge management literature. Section 3 discusses the background of the organization that is the source of this analysis, the FAO (Food and Agriculture Organization, an agency of the United Nations) and its technology section WAICENT (World Agriculture Information Centre), and discusses why FAO needs to have a multilingual capability. Section 4 reviews the basic nature of FAO's multilingual system. Section 5 provides an overview of some FAO multilingual knowledge resources, including a particular controlled vocabulary that is maintained by FAO in five different languages and some categorization schemes used by FAO and WAICENT. Section 6 provides a brief analysis of FAO and WAICENT's architecture, reviewing the important role of XML. Section 7 discusses the limitations of building multilingual systems. Section 8 provides a model to summarize our findings and analyzes. Finally, Section 9 briefly summarizes the paper and discusses some extensions.

2. Selected previous research in multilingual DSS and knowledge management

Multilingual systems have begun to find use in a large number of settings, including government, medical systems and libraries. In addition, some of the most important technical issues in multilingual systems are ontologies, since they help facilitate communication, structure and search about knowledge issues. Peters and

Sheridan [23] provide a brief history of the integration of multilingual resources into systems.

There has been limited research in multilingual DSS. For example, prior to 2006 there was a single paper in *Decision Support Systems* focused on multilingual systems [3]. That paper dealt with the issue of how to use group decision support systems in order to break down barriers to communication. Although there has been limited research in multilingual DSS, there has been research in knowledge management and multilingual systems, some of which is relevant to case study discussed in this paper.

2.1. *Knowledge management*

While providing appropriate knowledge “content,” some of the key functions of a knowledge management (KM) system include “converting” and “connecting” knowledge ([19,20] and others). Content includes a broad range of resources, such as knowledge about how to solve particular problems (“things gone right”) or knowledge about how to not solve problems (“things gone wrong”). Other critical content can include ontologies used to structure and search and to facilitate communication [20]. Individual knowledge needs to be converted to group available knowledge, and data and text need to be converted to usable knowledge, not to be lost in the piles and piles of data and text that are available. Although not discussed at the time of that research, in a multilingual system, knowledge resources also need to be converted from one language to others, as we see in the case study in this paper. As a result, the need for multilingual capabilities broadens the base and requirements of knowledge management in general.

Further, individuals need to be connected to knowledge resources and people. Knowledge resources need to be searchable and links between appropriate knowledge islands need to be established. In [19] and others, the need for consideration of multiple languages was not considered in the case of “connection.” Again multilingual requirements need to be considered since the connection needs to be among those of matching languages, otherwise the connections will not be helpful. Search needs to provide useful connections in the language(s) appropriate for the user. Further, connecting knowledge to other knowledge also must consider language, since in general, knowledge in different languages cannot be consumed by individual users. As a result, from a connection and conversion perspective, some characterizations of knowledge management have been historically underspecified. Section 8 presents a model based on these extensions and categorizes FAO and some other systems based on the system information discussed later in this paper.

2.2. Multilingual support in government

Multilingual systems can play an important role in government activity. As we noted above, governments may have multilingual consumers of knowledge resources.

Beyond general interaction with the populace, emerging areas of government multilingual interest include terrorism and crime analysis. For example, Last et al. [13] and Qin et al. [24] have investigated the use of multilingual approaches to discover the presence of terrorists groups on the Internet. Terrorists use multiple languages, so the systems that are used to find them need to consider and understand multiple languages. Similarly, crime does not limit itself to a single language. Thus, Yang and Li [30] discuss how to extract multilingual information for crime analysis focusing on Chinese and English documents.

2.3. Multilingual medical systems

Sevinc [26] and others have stressed the need for medical research to be available in multiple languages. Further, there have been some multilingual DSS developed for support of medical problems. For example, Ohmann et al. [21] developed a decision support system for the diagnosis of abdominal pain. The system involves participants in 18 different countries, primarily from Eastern Europe. One of the primary concerns of the system is the data dictionary which defines an ontology of medical terms. In addition, the system contains those data dictionary terms in multiple languages, including Polish, Romanian, Estonian and others. In another DSS, Goble et al. [6] create multilingual terminology server designed to provide an ontology to a broad range of medical applications. As another example, Zhou, Qin and Chen [31] focused on facilitating the search for Chinese medical information.

2.4. Personalized library portals

Since knowledge management systems provide access to multiple resources, one comparable source are libraries. Extending personalized portals such as “My Yahoo!,” there has been a sequence of research related to the development of personalized library portals. Starting with Morgan [17] and Cohen et al. [5] libraries have allowed users to create personal web pages to capture and store frequently used electronic library resources. There have been a number of updates to that original concept and views of the future (e.g., [4]). In addition, there have been multilingual views of the MyLibrary concept. For example, as noted by Rozic-Hristovski et al. [[25],

p. 157], “One of the most important needs of visitors from ... abroad is multilingual support, which means that the users can select a language in which the portal interface is presented to them.”

2.5. Ontologies in multilingual knowledge management systems

There are many design considerations associated with multilingual systems (e.g., Starr [27]). One of the most important design considerations is the design, development and use of ontologies.

Ontologies have been defined as explicit specifications of conceptualizations (Gruber [7]). Ontologies provide a shared vocabulary and a common language that can be used for many purposes, including indexing, search and retrieval (e.g., [20]). Ontologies facilitate reuse of knowledge resources. However, because they are based in language, there are a number of impediments to their usage, such as changing meaning over time (e.g., [18]). Unfortunately, in that analysis of impediments of ontologies, the multilingual nature of ontologies was not addressed. As a result, we can see limitations in previous knowledge management models because of a lack of consideration of multilingual factors. Additional factors come into play, as ontologies are likely to be generated in the particular language of concern (e.g., English). However, there may or may not be a particularly good match in other languages or there may be multiple terms, rather than a single term.

In addition, it is this later notion of multilingual for which ontologies are so important since they facilitate use of multiple languages. Mayer [16] provides some tools to facilitate terminological databases for multilingual ontologies. Jarrar et al. [8] discuss an ontology-based complaint system where they use the complaint ontology to capture the core knowledge in the domain and show how the multilingual complaint environment can be simplified using an ontology based approach. Vouros et al. [29] use an ontology-based approach to provide the basis of search and navigation of information in a multilingual knowledge management system.

2.6. Complexity of “multilingual” systems

As we have seen in this literature review, “multilingual” seems to have many gradations. As noted by Rozic-Hristovski et al. [25], at one extreme, multilingual means being able to select a web portal interface language. At the other end of the spectrum, not only the interface but also the resources are available in multiple languages. For example, Peters and Sheridan ([23],

p.52) refer to “...accessing, querying and retrieving information from collections in any language”

Further, the system or approach can be either active or passive in its conversion of multilingual resources. In a passive approach, the knowledge resources taken are the knowledge resources used. In an active approach, knowledge resources would need to be translated to the other languages. Resources in multiple languages need to be organized, indexed and searched, typically in multiple languages. As a result, movement to multiple languages increases the complexity of the knowledge management system.

3. Case study background

Throughout this paper we analyze key characteristics and emerging issues using the multilingual knowledge management system of the *Food and Agriculture Organization* (FAO) and their information group, *World Agriculture Information Centre* (WAICENT) [11]. These organizations were chosen for a number of reasons. First, because these are public organizations, there is information about these groups and it is publicly available. Since information is publicly available, it is not considered contrary to the organization to discuss this information. Second, these and other public organizations have many constituents, from different cultural and language backgrounds. Unlike multinational corporations, these organizations cannot dictate a particular language be used, and thus must have multilingual systems. Third, I have found that the organization is a good single source to discuss this wide range of issues.

3.1. FAO

Prior to the end of World War II, in 1943, 44 different countries came together in a meeting in Virginia, committing to an international organization for food and agriculture (<http://www.fao.org/docrep/meeting/010/j6285e/j6285e04.htm>). Ultimately, a report was issued in 1945 on behalf of the United Nations that resulted in the founding of FAO. Since its founding, it was agreed that FAO should be “concerned with that large sector represented by the world’s farms, forests, and fisheries, and by the needs of human beings for their products.”

Now, FAO (*Food and Agriculture Organization*) is an agency of the United Nations, headquartered in Rome, Italy. It has offices in more than 80 countries, with a staff of about 4000 people. FAO has its own governing body and it has a core budget of roughly \$ 640 million, with almost an equal amount of outside funding.

The purpose of FAO as captured in the first article of the FAO Constitution is that “The Organization shall collect, analyze, interpret and disseminate information relating to nutrition, food and agriculture.” FAO also makes recommendations on national and international action in a variety of arenas, such as scientific research in agriculture and improvement of processing, marketing and distribution of food and agricultural products. FAO’s reach is broader than just food. For example, one of their major concerns at the writing of this paper is the Avian Flu and its ramifications.

3.2. WAICENT

WAICENT’s (*World Agriculture Information Centre*) is FAO’s strategic program on information management dissemination (e.g., http://www.un.org.pk/library/unirr_waicent.html). WAICENT provides electronic access to FAO’s information resources through its portal. In particular, WAICENT makes FAO’s knowledge on all fields of food security and agricultural development widely available to users around the world, through the Internet as part of a large knowledge management system.

WAICENT’s scope includes organizing and linking information in order to facilitate user access; providing visualization systems; and providing decision-support systems at national levels in order to help achieve food security through use of information. In countries where there is limited Internet access, they provide high speed Internet connections to access their information.

3.3. “Digital inclusion” and “digital divide”

An increasingly important issue, particularly to organizations such as FAO is the notion of “digital inclusion” (e.g., Verdu et al. [28]). “Digital inclusion” relates to the fact that different countries and thus different cultures and languages have different access to digital technology and resources. Digital inclusion can be a driving force that influences an organization’s strategy and, ultimately, the extent to which they provide multilingual resources.

Perhaps, the most visible digital inclusion issue is that of the Internet and access to knowledge resources. Unfortunately, some countries do not have the level of access as others. For example, the so-called *Meda Countries* (Algeria, Egypt, Jordan, Palestine, Morocco and Turkey) generally have Internet penetration levels of less than 10% of their populations, resulting in a “digital divide.” Internet access, computers and even electricity can be stumbling blocks to “digital inclusion” in the use of knowledge management resources. In addition, the lack of

available multilingual knowledge resources can limit the push for use of resources from an Internet environment.

Although digital inclusion was not an issue at the time of FAO's founding, digital inclusion is an increasingly important issue and embedded in the rationale for multilingual systems. In order to try and limit differentials in access to digital resources, multilingual systems for infrequent users can be an issue that needs to be addressed.

3.4. Need for multilingual capabilities

FAO and WAICENT require a multilingual Internet presence for a number of reasons, including

- A single voice to many users
- FAO is a frequently visited web site
- FAO is visited by a broad base of users
- Their constitution basically requires it.

3.4.1. A single voice to many users

First, they represent over 180 different countries, thus suggesting a broad based set of users. Internally, FAO is a highly decentralized organization, with data from over 200 different groups, generated by different development groups within FAO. However, they need to show externally a single "voice to the world," but that voice needs to be expressed in the five official languages (English, French, Spanish, Chinese, and Arabic) as well as Russian and other local variations. Ideally, the same question has the same answer, independent of the language.

3.4.2. FAO is a frequently visited web site

Second, FAO is a very frequently visited site. Thus, multilingual capabilities can be leveraged to many users. FAO is a web site that receives hundreds of millions of visitors each year. Their site gets hundreds of millions of hits and millions of visits each year. (A hit is a request to a server for a file. A visit is a measure of the number of unique users who visited a web site during a certain time period.)

Table 1
User sessions

Source	Percent
North America	57
Europe	30
Latin America	5
Asia	4
South Pacific	2
Africa	2

Source: [9].

3.4.3. FAO is visited by a broad base of users

Third, FAO's web sites get visited by users from all over the world. The broad based use of FAO and WAICENT is apparent with an analysis of the user sessions. North America recorded the largest number of user sessions. However, user sessions were also captured from all over the world, as seen in Table 1. Accordingly, the need for multilingual knowledge management capabilities permeates WAICENT's Internet presence.

3.4.4. Their constitution basically requires it

As noted above, one of the key purposes of FAO is to collect, analyze, interpret and disseminate information. Since they represent many different countries, in order to perform those activities, FAO must build a multilingual capability.

4. FAO's multilingual Internet presence

How does FAO provide the information that they need to in order to meet their scope requirements? A number of their key efforts have multilingual aspects.

4.1. FAO home page

FAO's homepage provides a multilingual capacity of Arabic, Chinese, English, French, and Spanish. Fig. 1 contains an English version. More than just multiple languages are required. For example, in the Arabic page, the layout of information is different than in the English version. For example, when information is presented in Arabic, it is presented right to left, rather than left to right. However, with the "reversed" layout, the same icons are used, so that users likely get a page experience that is geared toward their specific language, but with many similarities to other languages. As a result, there is no "preference" given to any particular language.

4.2. WAICENT home page — portal

The WAICENT portal is more compact, and less "grabbing" (e.g., there are no photos, and the colors are subdued greens and blues) than the FAO page since it is more focused on presenting a broader base of resources. A screen print illustrating the home page is in Fig. 2.

The WAICENT portal provides links to over 250 sites. In addition, it has access to search engine "WAICENT Information Finder," to help users find the information for which they are looking. The FAO Web site is a large site, having approximately 500,000 web pages and over 100 databases [11].



FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS

helping to build a world without hunger

العربية 中文 français español

Search **NEW**

What is FAO?

From the Director-General

Governing Bodies

Agriculture, Biosecurity, Nutrition and Consumer Protection

Economics and Statistics

Fisheries

Forestry

Sustainable Development

Technical Cooperation

Programme and Budget; Evaluation

Legal Office

Decentralized Offices

Employment

Procurement Service

Interdisciplinary Activities: Trade, Biotechnology, Gender...

Statistical Databases

Country information

Virtual Library

Publications and Documents

more...

International Partners

UN system Network on Rural Development and Food Security

Food Insecurity and Vulnerability Information and Mapping Systems

NGOs and civil society organizations

FAO and the Private Sector

Working together with IFAD and WFP

International Alliance Against Hunger

Other UN Organizations

Newsroom

FAO's Regional Conference for Africa opens
Experts and officials from some 53 countries meet in Mali as food shortages continue in many African countries

FAO appeals for \$40 million for agricultural relief and recovery activities in Sudan
Humanitarian assistance coupled with longer-term development aid crucial to lasting peace

Microfinance crucial to alleviating poverty in forest communities
Basic financial services can help families start their own businesses

Bird flu virus could spill over to Africa and Europe in springtime
Fighting the disease in animals is crucial to win the battle against the virus

[more news...](#)



FAO is helping returnees in Sudan rebuild their livelihoods in livestock-raising, agriculture and fisheries

[more photos...](#)

Focus on the Issues

Communities reborn: much progress visible in year after tsunami

Small Island Developing States struggle to survive

FAO Reform - A vision for the 21st century

Millennium Development Goals

FAO: Serving its members

Meetings

FAO Regional Conferences 2006

112th Finance Committee
Rome, 9 February 2006
FAO Council

Knowledge Forum

World Agricultural Information Centre

Avian Influenza

Tsunami

Progress since the World Food Summit

World Food Day

FAO Hunger Map

TeleFood

Feeding Minds Fighting Hunger

Selected Key Programmes

SPFS: Special Programme for Food Security

EMPRES: Emergency Prevention System for Transboundary Animal and Plant Pests and Diseases

GIEWS: Global Information and Early Warning System on Food and Agriculture

TCE: Emergency Relief and Rehabilitation

Current Special Initiatives

Current Desert Locust Situation

Small Island Developing States

Spearheading Regional Food Security

FAO Trust Fund for Food Security

Fig. 1. FAO homepage (English).

WAICENT has enabled the portal home page to be available in five languages: Arabic, Chinese, English, French and Spanish. Simply by clicking the language in the blue tool bar enables the user to choose the language that they prefer to use.

4.3. Available resources

FAO provides their users with a number of different types of information and knowledge resources, including web pages, press releases, pictures, documents,

WAICENT PORTAL

Home WAICENT Information Finder Ask FAO

العربية 中文 English Français Español

WAICENT Highlights

20 January 2006 Now in Spanish: IMARK E-learning Module "Management of Electronic Documents"

9 June 2005 Making a mark in information management - FAO releases training package on creating digital libraries

20 May 2005 New EasyPol Web site launched

3 February 2005 The AGORA initiative celebrates its first anniversary.

[More](#)

What's new on the FAO Web site

Glossary of FAO Databases and Information Systems

FAO Web site directory

This Web site directory provides links to approximately 250 main sites and subject entry points.

Animal Production & Health
[Animal Health](#), [Animal diseases & Control](#), [Animal genetic resources & Breeding](#), [Animal nutrition & Feeds](#), [Production & Livestock](#), [more](#)

Economics & Policy
[Agroindustry](#), [Economic development](#), [Policies](#), [Trade](#), [Marketing & Commodities](#), [more](#)

Education & Extension
[Communication for development](#), [Education](#), [Extension](#), [Training & Capacity building](#), [more](#)

Engineering, Technology & Research
[Biotechnology](#), [Geographical information systems](#), [Postharvest technology](#), [Research](#), [Statistics](#), [more](#)

Farming Practices & Systems
[Farm management](#), [Farming systems](#), [Land use](#), [more](#)

Fisheries & Aquaculture
[Aquaculture](#), [Development](#), [Ecosystems](#), [Governance](#), [Issues](#), [more](#)

Food Security
[Agricultural situation & Early warning](#), [Emergency relief & Food aid](#), [Ethics](#), [Food supply](#), [International cooperation](#), [more](#)

Food safety & Human nutrition
[Diet & Nutrition](#), [Food additives](#), [Food composition](#), [Nutrition education](#), [Quality controls & Assurance](#), [more](#)

Forestry
[Assessment & Monitoring](#), [Environment](#), [Forest management](#), [Forest products & Services](#), [Forest resources](#), [more](#)

Geographical & Regional Information
[Africa](#), [America](#), [Asia](#), [Europe](#), [Oceania](#), [more](#)

Government, Administration & Legislation
[Administration](#), [Agricultural and rural legislation](#), [Environmental legislation](#), [Fisheries](#), [Food legislation](#), [more](#)

Information Management
[Databases & Information systems](#), [Documentation](#), [Early warning systems](#), [Geographical information systems](#), [Information science](#), [more](#)

Natural Resources & Environment
[Biodiversity](#), [Climate change](#), [Desertification](#), [Drainage & Irrigation](#), [Ecology & Ecosystems](#), [more](#)

Plant Production & Protection
[Crops & Crop management](#), [Fertilizers](#), [Integrated pest management](#), [Irrigation](#), [Pest control & Pesticides](#), [more](#)

Rural & Social Development
[Gender](#), [HIV/AIDS](#), [Households](#), [Participation](#), [Poverty](#), [more](#)

Featured sites

- Agricultural Information Management Standards
- World Food Day - Agriculture and intercultural dialogue
- UN System Network on Rural Development and Food Security
- FAO World Reviews "FAO State of... Flagship publications"
- Feeding Minds, Fighting Hunger
- Special Programme for Food Security (SPFS)
- World Food Summit: *five years later*
- World Summit on Sustainable Development 2002

Fig. 2. WAICENT portal.

databases and controlled vocabularies. On the Internet these items need to be searched and indexed. Multilingual aspects of some of these resources are discussed further below in Section 5.

4.4. Search (WAICENT Information Finder)

Search across the FAO and WAICENT knowledge management resources is done using a number of mechanisms. However, the most prominent approach is the WAICENT Information Finder [2]. Information Finder allows for web-based information to be located using a number of approaches, including, free text, indexed text,

subject categories, type of resource and FAO organization unit [10]. That information finder uses the open source search engine "Nutch." Nutch uses a similarity metric to rank hits satisfying a query, returning a similarity score with each hit.

Search needs to be provided across multiple languages. As a result, the search page (Fig. 3) is available in five different languages. However, the results of a search are not limited to the findings in that language. As seen in Fig. 3, a search for "Katz" finds Chinese, Arabic and English citations in the first ten that it finds, from a French language search. Accordingly, search findings are integrated across multilingual documents. Further, the same

The screenshot shows the FAO WAICENT Information Finder interface. At the top, the FAO logo and the text 'FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS' are visible, along with the tagline 'helping to build a world without hunger'. Below this is a navigation bar with 'Home', 'FAO Home', 'WAICENT Portal', 'Ask FAO', and language options: 'العربية', '中文', 'English', 'Français', 'Español'. A search bar contains the text 'Katz' and a 'Search' button. Below the search bar, it indicates 'Hits 1-10 (out of about 282 total matching pages) in 0.01 seconds:'. A list of search results follows, each with a title, a brief description, a URL, and a 'cached' status. The results include: 'موقع معايير إدارة المعلومات الزراعية على الويب', '农业信息管理标准网站', 'موقع معايير إدارة المعلومات الزراعية على الويب', 'Expert Consultation on the Redesign of AGRIS and CARIS', 'Site web GIL', 'FAO 21st ERC REPORT: APPENDIX A', 'Current and Emerging Issues for Economic Analysis and Policy Research (CUREMIS II) - Volume I: Latin OF LATIN AMERICA - ELIZABETH KATZ Introduction Is there a ...', 'A Preliminary Study of Hydrogen Sulphide Toxicity on Juveniles of Macrobrachium rosenbergii (de Man) ... authors. Doudoroff and Katz (1950) reported that the ...', 'Agricultural Ontology Service Project (AOS)', and 'SD: Institutions : Technical Consultation on Decentralization. 16-18 December 1997.'.

Fig. 3. Search results in WAICENT.

search for “Katz” in each of the five available languages found the same number of items returned each time. Thus the same “picture” (or view) is painted for each user, independent of language. The system might be extended by allowing the user to choose the language in which the search results are returned, rather than returning results in arbitrary languages.

Multilingual search capabilities could be treated differently. For example, documents listed as “found” could be constrained to those of the language requested. However, many users are not interested only in documents in a certain language, but instead are interested in documents across all languages. Further, unfortunately, with that approach, different “pictures” would be painted for users with different languages. Potentially some documents, etc. would not be included with the different languages.

4.5. Keyword search (“Browse by Topic”)

Keyword search is also available on the FAO and WAICENT sites. On the WAICENT search engine, there is a choice to “Browse by Topic.” That choice provides the user with the ability to browse through a number of keywords, paging through each letter of the alphabet. As with the other pages, multiple languages are allowed.

An example keyword search is presented in Fig. 4. Continuing the search example, “Katz” is not among the keywords beginning with “K.” As a result, the keyword searches are using different information than Information Finder searches.

Unfortunately, with keywords, it is difficult to get the same “view” or “picture” for users across different

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS
helping to build a world without hunger

WAICENT INFORMATION FINDER

Home **FAO Home** WAICENT Portal Ask FAO العربية 中文 English Français Español

Free text Specialized search Browse by topic About Help

Browse by topic provides easy access to FAO documents (taken from the Corporate Document Repository) and FAO main sites by listing predefined keywords from A to Z. Click on a letter to search for terms or fill the box to search for a specific term.

A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z

Search for terms

Kalimantan [2]	Kerala [3]	Knowledge based systems [9]
Katsuwonus [1]	Kernels [3]	Knowledge management [3]
Katsuwonus pelamis [1]	Kerria lacca [1]	Korea Democratic People's Republic [21]
Kazakhstan [10]	King crabs [2]	Korea Republic [20]
Keeping quality [18]	Kingdom of Morocco [1]	Krill [2]
Kefir [1]	Kingdom of Nepal [1]	Kuwait [4]
Kenaf [1]	Kiribati [2]	Kyrgyzstan [4]
Kenya [62]	Knowledge and information systems [26]	

print-friendly version Comments? Please write to the Webmaster | © FAO 2005

Fig. 4. Keyword search (“Browsing by Topic”).

languages, for a number of reasons. First, as discussed below, keywords are typically provided by human indexers. Words are chosen for a document and then

translated. This can result in keywords that do not “fit” the document equally well in each language. Second, words in one language do not necessarily begin with the

EMPRES
 Emergency Prevention System
 for Transboundary Animal and Plant Pests and Diseases

Français Español العربية

Agricultural pests and diseases often migrate or spread across borders and cause major losses and emergencies. In the past, such damage has on occasions been catastrophic, leading to famines and sometimes triggering trade restrictions. Developing countries are frequently not able to react sufficiently quickly to such events, and extensive emergency operations as well as international assistance becomes necessary. Although effective control methods usually exist now against these pests and diseases, such crisis management inevitably involves delays, a low efficiency/cost ratio and an inability to contain the problem at an early stage. In 1994 FAO established an Emergency Prevention System (EMPRES) for Transboundary Animal and Plant Pests and Diseases in order to minimize the risk of such emergencies developing. Initial priority was given to two transboundary pest and diseases problems:

Animal Disease Component



Major transboundary livestock diseases, including Rinderpest and other epidemic animal diseases (contagious bovine pleuropneumonia, foot-and-mouth disease, contagious caprine pleuropneumonia, peste de petit ruminants, rift valley fever, and lumpy skin disease). These diseases are among the most contagious and place a serious burden on the economies of the countries in which they occur.

Desert Locust Component



The Desert Locust is an international problem due to the frequent migration of swarms across borders. Since earliest recorded history, this pest has been considered a serious threat to agricultural production in Africa, the Near East and Southwest Asia and often requires large-scale control operations.

FAO Home Page

Fig. 5. EMPRES.

same letter in another language. As a result, the paging through particular letters of the alphabet in different languages will not yield the same set of concepts. Accordingly, it is difficult to get the same view for each language.

4.6. User surveys

Periodically, WAICENT needs to understand how well the users think that the knowledge management system is working and to seek feedback. As a result, they have done Internet-based surveys of their users. This has been done by developing and circulating multilingual versions of the same survey.

Even user surveys also must be considered in multiple languages. For example, one example that was reviewed indicated that those completing the survey will reap one of three different benefits in the form of FAO gifts. That statement of available gifts and the survey also was available in multiple languages of Arabic, English, French and Spanish.

5. Multilingual knowledge resources

Not all resources are available in all five languages, all of the time. As two examples, “Emergency Alerts,” and “E-Learning” are available in fewer different languages. Resources and countries affected by the information influence those countries for which languages materials are developed. However, others are available in more than five, such as AGROVOC.

5.1. Emergency alerts (Emergency Prevention System)

In 1994, FAO established an *Emergency Prevention System (EMPRES)* for problems that cut across multiple countries boundaries, to try to minimize the risk associated with those emergencies. Current concerns include the “Avian Influenza,” and “Foot and Mouth Disease.” Currently, as seen in Fig. 5 there are alerts issued regarding particular world events that can impact the food supply in English, French, Spanish and Arabic.

The screenshot shows the FAO Agricultural Information Management Standards website. The main heading is "FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS" with the tagline "helping to build a world without hunger". Below this is the "Agricultural Information Management Standards" section, which includes a search bar and a navigation menu with options like Home, Partners, Discussion lists, etc. The main content area is titled "Browse the AGROVOC thesaurus" and shows a search interface with a search term field, search button, and radio buttons for search criteria (starting with, containing text, exact match). Below the search interface is a grid of letters for navigation and a dropdown menu for language (set to English) and records per page (set to 10). The results section shows "Terms found: 3396" and a list of terms with their termcodes and statuses.

Termcode	Term	Status
40070	Aaptosyax grypus	Descriptor with relations
3	ABA	Descriptor with relations
4	Abaca	Descriptor with relations
45978	Aballistes stellaris	Descriptor with relations
5	Abalones	Descriptor with relations
33507	Abamectin	Non-Descriptor with USE relation
6	Abandoned land	Descriptor with relations
7	Abattoir byproducts	Descriptor with relations
8	Abattoirs	Descriptor with relations
40071	Abbottina rivularis	Descriptor with relations

Fig. 6. Ontology in English.

However, the drill down associated with those alerts is not provided in all of the languages apparently unless they are necessary. For example, in a branch off of that page, “Locust Watch,” the languages of Arabic, French and English are presented. Spanish was not included, since apparently the Locust Watch did not affect any Spanish speaking countries in the situation update.

5.2. E-Learning

FAO is partnering with another group to generate some E-Learning capabilities (<http://www.imarkgroup.org/>). For example, as can be seen in Fig. 2 earlier, one of the E-Learning modules became available in Spanish on January 20, 2006. Currently, those resources are being generated in three languages, English, French and Spanish. As new

resources are developed they are released, whether or not they are available in all languages.

5.3. Controlled vocabularies at FAO: AGROVOC

As noted on the FAO web site, “AGROVOC is a multilingual, structured and controlled vocabulary designed to cover the terminology of all subject fields in agriculture, forestry, fisheries, food and related domains.” Lauser et al. [14] discuss some of the issues involved with development of an ontology that is ultimately used in a multilingual environment, focusing on an ontology used for biosecurity, in particular, a “Food Safety Ontology.” The process first developed the ontology in English. Then the ontology was translated manually into other languages.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS
helping to build a world without hunger

Agricultural Information Management Standards
 ■ Interoperability, Reusability, and Cooperation ■

Home | Partners | Discussion lists | العربية | 中文 | English | Français | Español

AGROVOC Concept Server | Metadata schemas | Tools | Publications | News/Events

AGROVOC Thesaurus

- Browse
- Sub-vocabularies
- Latest updates
- Suggest terms
- Download
- Webservices
- Copyright information

Knowledge Organization Systems

- By Type
- By Subject area

AOS/CS

- The Concept Server
- Applied ontologies in FAO
- Ontology relationships

Glossary

Frequently Asked Questions

Browse the AGROVOC thesaurus Last Update: November 2005

Search term: Search

starting with containing text exact match

A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z | All

Language: Chinese Show records per page: 100

Terms found: 36399

Pages: 1 2 3 4 Next >> Last

Termcode	Term	Status
3	脱落酸	Descriptor with relations
4	蕉麻	Descriptor with relations
5	鲍属	Descriptor with relations
6	废弃地	Descriptor with relations
7	屠宰场副产品	Descriptor with relations
8	屠宰场	Descriptor with relations
9	腹部	Descriptor with relations
10	冷杉属	Descriptor with relations
11	欧洲冷杉	Descriptor with relations
12	香脂冷杉	Descriptor with relations
13	科罗拉多冷杉	Descriptor with relations
14	北美冷杉	Descriptor with relations
15	落矶山冷杉	Descriptor with relations
16	高加索冷杉	Descriptor with relations
17	高大冷杉	Descriptor with relations
18	萨哈林冷杉	Descriptor with relations
19	西伯利亚冷杉	Descriptor with relations
20	异常行为	Descriptor with relations

Fig. 7. Ontology in Chinese.

Multilingual domain ontologies, also referred to by FAO as “Controlled Vocabularies,” facilitate knowledge disbursement and communication [1]. They can be used to index documents, etc. to facilitate search.

“AGROVOC” illustrates some of the issues associated with multilingual knowledge management sites. As seen in Fig. 6 there are five different languages (Arabic, Chinese, English, French and Spanish) for which the web page description “around” the ontology is available. However, there are fourteen different language versions of the ontology, either on the FAO site, other sites or in process.

Fig. 6 presents a screen shot of the ontology in English. On that page, the description on the page “around” the ontology is in English, the ontology is in English and the description within the ontology is in English. Similarly, the same ontology is available in Spanish, with all three elements also in Spanish. However, for some languages, although the description around the ontology is in that language, and the ontology is in that language, the descriptor within the ontology is in English. Further that the alphabetic order in each case corresponds to the appropriate ordering for the particular language and does not follow the order in English. Fig. 7 presents the Agrovoc page for the Chinese version of the ontology, however, the ontology is not in Chinese. The Chinese version of the ontology has the description around the ontology in English. Note that for this version, the web page around the ontology is in English, the descriptors are in English.

The progression in these exhibits illustrates a “life cycle” associated with multilingual systems. First, the page “around” the ontology is easier to change to different languages, than the ontology or the description of the ontology. Second, the translated ontology itself is more likely to be ready sooner than the documentation supporting the ontology. Third, in some cases there is no equivalent concept for the ontology in the target language. In those cases, where there is no concept match, the

English version is used. Fourth, the documentation in support of the ontology is likely to be the last developed. As a result, we see the documentation in English in some of the translations.

5.3.1. Number of items in AGROVOC for different languages

The number of items in each version of the ontology apparently is different as seen in Table 2. There is not a one-to-one translation between the versions of the ontology. There are a number of potential reasons for this. First, differences between the languages could result in differences in ontologies. A word in one language does not necessarily have a direct counterpart in some other language. Second, the translations were made by people. The quality of the translator can influence the extent to which there is the appropriate set of matches between ontology versions in different languages. Third, since the translations are made by people there can easily be errors in the translated ontologies.

A translated version of a controlled vocabulary can be analogous to an island in terms of Darwin evolution. If the vocabulary is left to develop on its own in that language, it is likely to take on changes that are different than the same ontology in another language also left to develop on its own. Accordingly, there needs to be a coordinated effort to ensure an “equivalent” similar view is provided in each language.

5.3.2. Adding to AGROVOC's multilingual capabilities

Two features of AGROVOC can provide increased multilingual capabilities (e.g., [15]). First, as seen in Fig. 8, there is the ability to suggest new terms that should be added to AGROVOC. As part of that addition feature, the user can recommend a term or concept in multiple different languages.

Second, additional languages can be added to AGROVOC with communication with FAO, as seen in Fig. 9, at the bottom. An excerpt from an AGROVOC page provides contact information for a potential contributor to provide vocabulary information to FAO in a number of subject areas, including (not in this screen shot) agriculture, geography and history, administration and legislation, economics, and sociology, and other areas.

5.3.3. Categorization schemes: AGRIS and CARIS

FAO's WAICENT also employs multilingual categorization schemes that facilitate search through sets of documents. One such scheme is “AGRIS.” AGRIS is the international information system, created by FAO in 1974, for the agricultural sciences and technology. AGRIS was designed to facilitate

Table 2
Number of items in different ontology versions

Version	Number of items
Arabic	25,883
Chinese	36,399
Czech	38,663
English	39,095
French	38,261
Japanese	38,651
Portuguese	36,325
Spanish	41,534
Thai	25,411

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS
helping to build a world without hunger

Agricultural Information Management Standards
 ■ *Interoperability, Reusability, and Cooperation* ■

Home | Partners | Discussion lists | العربية | 中文 | English | Français | Español

AGROVOC Concept Server | Metadata schemas | Tools | Publications | News/Events

■ **AGROVOC Thesaurus**

- Browse
- Sub-vocabularies
- Latest updates
- **Suggest terms**
- Download
- Webservices
- Copyright information

■ **Knowledge Organization Systems**

- By Type
- By Subject area

■ **AOS/CS**

- The Concept Server
- Applied ontologies in FAO
- Ontology relationships

■ **Glossary**

■ **Frequently Asked Questions**

Suggest terms

Thank you for your interest in sending us proposals for terms you would like to see included in AGROVOC.

Please send your proposals by email to: fao-agris-caris@fao.org or use the following submission form. When sending the request please specify the following:(fields marked with (*) are mandatory)

Term in English:

Term in French:

Term in Spanish:

Term in Arabic:

Term in Chinese:

Status of the term:

Reason for the proposal:

Additional Comments:

Organization Name:

Contact Name:

Contact Email: *

Thank you,
 AGROVOC Team, FAO.

© FAO, 2006 | Comments?Please write to the webmaster print-friendly version

Fig. 8. Suggest terms.

Human nutrition

- Human nutrition - General aspects
- Physiology of human nutrition
- Diet and diet-related diseases
- Nutrition programmes

Pollution

- Pollution
- Occupational diseases and hazards

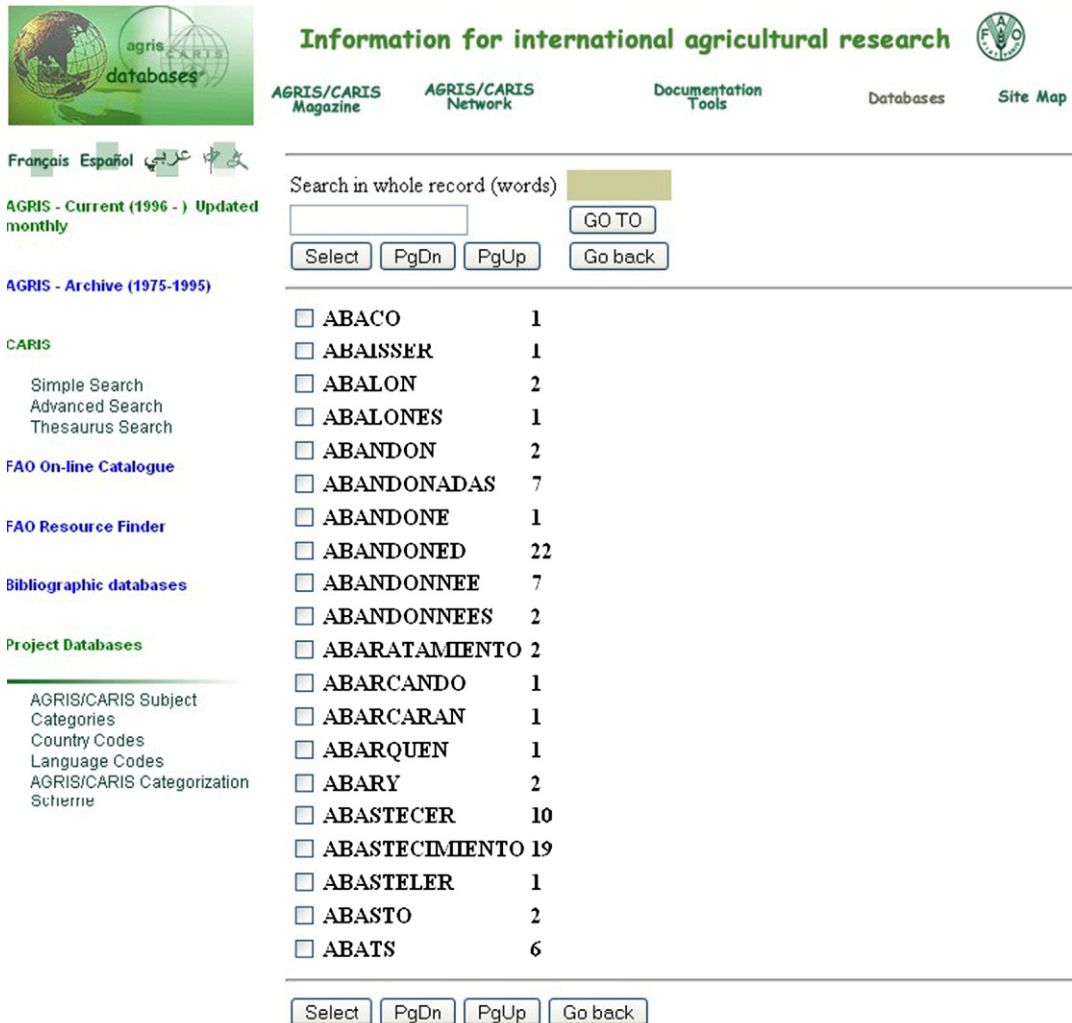
Auxiliary disciplines

- Mathematical and statistical methods
- Research methods
- Surveying methods

Do you want to contribute with another language? [Contact us](#)

© FAO, 2006 | Comments?Please write to the webmaster print-friendly version

Fig. 9. Excerpt from AGROVOC page “Contribute with Another Language”.



The screenshot shows the AGRIS and CARIS website interface. At the top, there is a logo for 'agris CARIS databases' and the text 'Information for international agricultural research'. Below the logo, there are navigation links for 'AGRIS/CARIS Magazine', 'AGRIS/CARIS Network', 'Documentation Tools', 'Databases', and 'Site Map'. On the left side, there are language options: 'Français', 'Español', 'عربي', and '中文'. Below the language options, there are links for 'AGRIS - Current (1996 -) Updated monthly' and 'AGRIS - Archive (1975-1995)'. Under 'AGRIS - Archive (1975-1995)', there are links for 'CARIS', 'Simple Search', 'Advanced Search', and 'Thesaurus Search'. Below these, there are links for 'FAO On-line Catalogue', 'FAO Resource Finder', 'Bibliographic databases', and 'Project Databases'. Under 'Project Databases', there are links for 'AGRIS/CARIS Subject Categories', 'Country Codes', 'Language Codes', and 'AGRIS/CARIS Categorization Scheme'. The main content area shows a search interface with a text input field, a 'GO TO' button, and 'Select', 'PgDn', 'PgUp', and 'Go back' buttons. Below the search interface, there is a list of keywords with checkboxes and counts:

<input type="checkbox"/>	ABACO	1
<input type="checkbox"/>	ABAISSE	1
<input type="checkbox"/>	ABALON	2
<input type="checkbox"/>	ABALONES	1
<input type="checkbox"/>	ABANDON	2
<input type="checkbox"/>	ABANDONADAS	7
<input type="checkbox"/>	ABANDONE	1
<input type="checkbox"/>	ABANDONED	22
<input type="checkbox"/>	ABANDONNEE	7
<input type="checkbox"/>	ABANDONNEES	2
<input type="checkbox"/>	ABARATAMIENTO	2
<input type="checkbox"/>	ABARCANDO	1
<input type="checkbox"/>	ABARCARAN	1
<input type="checkbox"/>	ABARQUEN	1
<input type="checkbox"/>	ABARY	2
<input type="checkbox"/>	ABASTECER	10
<input type="checkbox"/>	ABASTECIMIENTO	19
<input type="checkbox"/>	ABASTELER	1
<input type="checkbox"/>	ABASTO	2
<input type="checkbox"/>	ABATS	6

At the bottom of the list, there are 'Select', 'PgDn', 'PgUp', and 'Go back' buttons.

Fig. 10. AGRIS and CARIS partial key word list.

information exchange, in order to bring together the disparate literature dealing with the many different aspects of agriculture. AGRIS is a system in which the different participating countries are responsible for inputting references to the literature they produce. Then the 240 different participants can draw on all the information placed in the system.

Fig. 10 illustrates some of the multilingual capabilities developed to-date for “AGRIS.” The basic home page for AGRIS can be seen in English, French or Spanish.

5.3.4. Search in AGRIS

As seen in Fig. 10, AGRIS supports multilingual search. AGRIS searches for requested keywords in documents. Further descriptors in English, French and Spanish

can be used to find documents. Further, the user can then choose which languages in which the information is found can be displayed.

However, in the current system, the search does not support accented letters. As an example from the FAO web page, “When searching for terms, please enter the letter without the accent, e.g. to search for the Spanish word ‘maiz,’ you must enter ‘maiz’.”

5.3.5. CARIS

A closely related database to AGRIS is the Current Agricultural Research Information System (CARIS). Within CARIS keyword search includes a single list of keywords, rather than one list of keywords for each language, although either approach could be used. As an

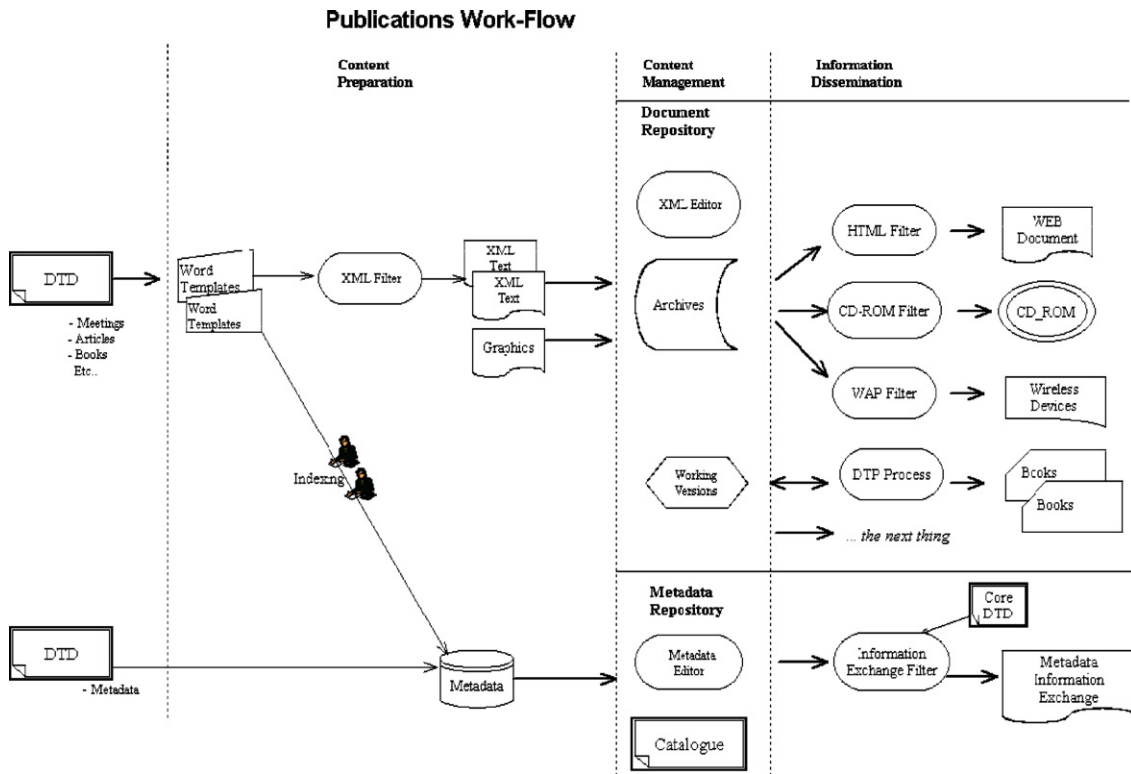


Fig. 11. Publication work flow.

example, Fig. 10 provides a page of keywords that includes keywords from multiple languages.

6. WAICENT multilingual work flow and architecture

In order to facilitate multiple languages, WAICENT employs a web services model using XML to exchange data in a multilingual environment, using both SOAP and RDF [11].

6.1. Work flow

WAICENT is a large scale operation that produces and adds many digital documents daily. As part of the infrastructure to support the production of documents for WAICENT, work flow is created. At the most basic level, documents are created, indexed by people and added to the repository. The documents are provided in multiple languages or are translated to multiple languages. Accordingly, the indexing and editorial activity needs to accommodate the multiple languages. A summary of the FAO and WAICENT work flow is given in Fig. 11 [e.g., 10 and 11]. Document Type Definitions (DTD) in the form of meeting proceedings, books, etc., in Word format are changed to XML (Extensible Mark-up Language) and

indexed using keywords by human indexers. That index information provides meta data about the particular documents. The data is then managed in its XML format and disseminated using a number of different environments, including HTML.

6.2. Architecture

FAO and WAICENT have adopted an architecture based on SOAP (Simple Object Access Protocol), RDF (Resource Description Framework) and XML, as seen in Fig. 12 (e.g., [11]). RDF provides the grammar (RDF Schema) and the syntax (RDF) to define declarative relations.

From a multilingual perspective, in the cached XML ontology, the variable “lang” is used to mark whether the data is in “EN” (English), “ES” (Spanish) or “FR” (French).

AGRIS’s search engine is based on XML-based as seen in Fig. 13 (e.g., [10]). Databases of documents and other materials are indexed using XML so that search engines can perform XML-based queries. This approach allows the simultaneous search across multiple databases on the Internet in real-time, yet also presents the user with a single set of results [15].

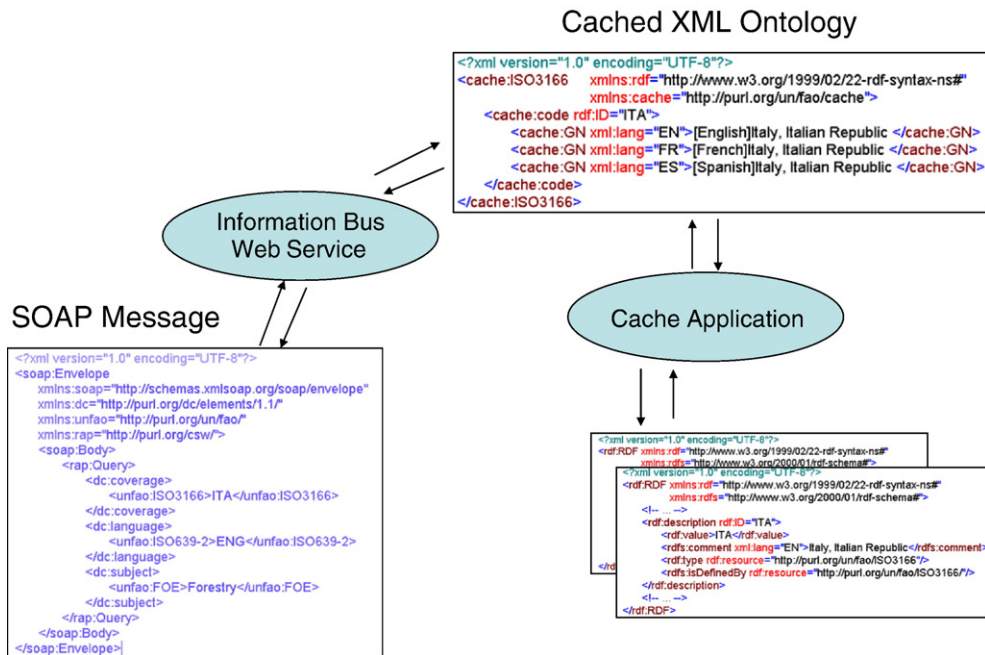


Fig. 12. Architecture.

Although XML provides the technology for vocabularies and ontologies, until recently, search and other capabilities were limited because various countries that provided content did not use the same structures. As a result, recently, AGRIS has adopted an “exchange layer” approach to the problem of having multiple document standards and to avoid requiring that information be put in a single format (Onyancha et al. [22]). Using this approach, FAO recently made the entire AGRIS repository of over 3 million records, available in the specifically developed “AGRIS AP Exchange Format.” This approach is based on using agreed on meta data to facilitate data exchange [22]. However, the primary difficulty associated with getting this exchange layer to work is getting stakeholders to agree on the underlying ontologies and vocabularies (e.g., Keizer [12]). If ontology development is difficult in a single language environment [18], those difficulties are compounded in a multilingual and multi cultural environment.

6.3. Implementation of web services in AGROVOC

Web services have been implemented across FAO and WAICENT applications. Web services provide an efficient way to approach providing multilingual capabilities. Particular web services can be designed to meet specific multilingual capabilities. A summary of some of the web services available are provided in Fig. 14 (e.g., [15]).

Although page information is provided in multiple languages, it is interesting to note from a multilingual perspective that the web services names are in English. Accordingly, the names themselves can loose context when put into other language settings, such as the one in Fig. 14, depending on the user. Alternatively, web services *names* could also be translated, but refer to the same web service through the appropriate links or other approaches. If they need a new language, they would not necessarily need a new site, they typically could just need another column as part of a table look-up.

7. Multilingual knowledge management systems: FAO and WAICENT

This case study has brought out some issues related to multilingual knowledge management over the Internet. The purpose of this section is summarize and integrate the findings of this paper.

First, as we have seen and can anticipate, such multilingual development requires substantial resources to both develop and ensure that the translation of a document is appropriate. As a result, development of multilingual knowledge management capabilities can be costly and time consuming because of the replication of knowledge and the complexity added to the system from the need to be multilingual. Development of a multilingual capability requires establishing workflows and integrating appropriate technologies as seen in Fig. 11. Although

AGRIS Multiple Database Search Engine

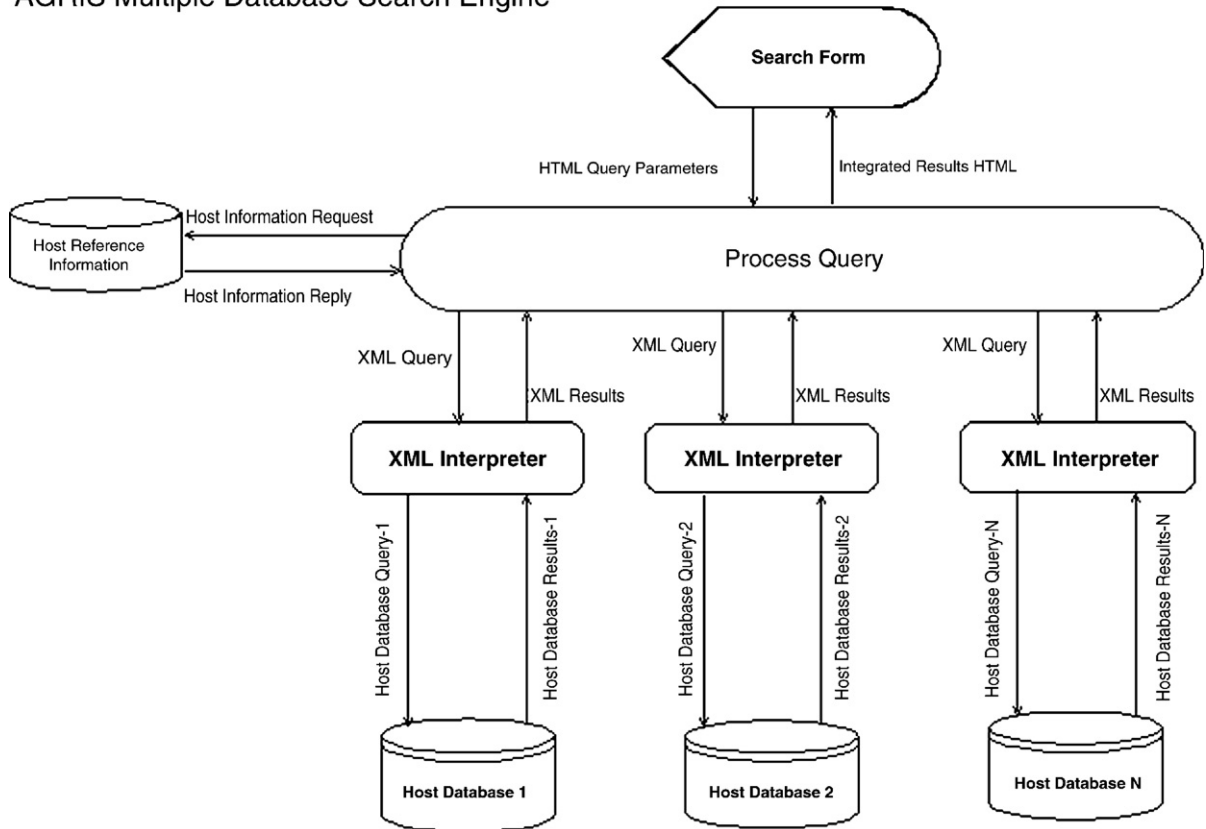


Fig. 13. AGRIS XML-based search.

outside of the scope of this paper, it potentially means reengineering and redesigning flows to accommodate new technologies over time.

Second, different language versions of documents are not all available simultaneously. As a result, translations are published on the web as they are developed, rather than waiting for all translations of a given document to be completed, before allowing any to be published. Unfortunately, this can lead to the appearance (but not fact) of apparent development priorities. Strategies need to be established to determine resource allocations. If a plague appears headed for a particular area of the world, resources need to be made available to facilitate that area, potentially driving the decision as to multilingual development sequence.

Third, when resources are developed in one language, but not another, it is not clear how that information should be communicated to the users. One approach is to let the user choose a “back-up” language where the resources are available. Then resources are listed in the primary and back-up language. Another approach is to allow for machine translation, which unfortunately, is often not a

good alternative. This issue also leads to an important search issue.

Fourth, in the case of searching multilingual resources, there are a number of potential choice issues. Allowing the user to choose an interface language is critical to letting them initiate a search. However, it may also be important to allow the user to target particular languages in their search, or be able to translate resources found in other languages.

Fifth, some of the greatest poverty exists in Africa. Unfortunately, Africa is home to hundreds of languages. The sheer number of languages makes catering to each of the individual languages virtually impossible. However, issues such as the digital divide and providing digital resources to problem areas may drive development of digital resources to a relatively small number of users.

Sixth, although information can be critical to alleviating hunger and malnutrition, ultimately, distribution of information over the Internet is limited to those locations with basic infrastructure, including electricity. Thus, it is important in those settings for FAO and WAICENT to provide the facilities for users in such settings to access FAO materials. This is seen in the provision of a number of



Normes pour la gestion d'information agricole

■ interopérabilité, réutilisabilité et coopération ■

 Recherche

[Accueil](#) | [Partenaires](#) | [Listes de discussion](#) | [العربية](#) | [中文](#) | [English](#) | [Français](#) | [Español](#)
[Serveur du concept AGROVOC](#) | [Schémas de métadonnées](#) | [Outils](#) | [Publications](#) |

Actualités/Événements

■ Thésaurus AGROVOC

- Naviguer
- Sous-vocabulaires
- Dernières modifications
- Proposer un nouveau terme
- Télécharger
- Services en ligne
- Droits d'auteur

■ Systèmes d'organisation des connaissances

- Par type
- Par domaine

■ SOA/SC

- Serveur de concepts
- Ontologies appliquées de la FAO
- Relations d'ontologies

■ Glossaire

■ Questions fréquemment posées

Services web AGROVOC

Il est aujourd'hui possible d'accéder à AGROVOC via les services web qui peuvent être appelés à partir de n'importe quelle application client. Les services web sont assurés par Apache Axis qui utilise le logiciel Tomcat. Ces services sont appelés par le biais d'appels SOAP classiques et renvoient une réponse SOAP standard. Pour de plus amples renseignements sur l'utilisation de la technologie utilisée, consultez les sections [téléchargements](#) et [liens](#) ci-dessous.

Services web existants

Cliquez sur l'un des services web suivant pour faire un essai en ligne!

(Il est possible de télécharger la description complète des services web dans la section Téléchargement ci-dessous)

Nom du service web	Description
getTermcodeByTerm	Renvoie le code-terme d'un terme AGROVOC existant
getTermByLanguage	Renvoie le terme AGROVOC spécifié par un code-terme et une langue
getAllLabelsByTermcode	Renvoie toutes les étiquettes d'un terme correspondant au terme-code spécifié
searchByTerm	Renvoie tous les termes AGROVOC contenant la chaîne de recherche spécifiée
getConceptByTerm	Renvoie le concept, c-à-d les étiquettes, les termes plus larges, les termes plus étroits, les termes associés correspondant à un terme-code
getDefinitions	Renvoie une définition du terme, un historique ou des notes de contenu

Téléchargements

Document	Description	HTML	Télécharger
WSDL	Description WSDL des services web	fichier xml	-
Documentation des services web	Documentation détaillée des services web existants	-	word doc (56KB)
	C'est une simple application JSP (Java Server Pages) pouvant être facilement installée dans le		zip file

Fig. 14. AGROVOC web services.

Internet — supported facilities around the world that they provide in order to try and mitigate the digital divide.

Seventh, developers of multilingual systems must address the issue of connecting multilingual resources. Do they capture and present related knowledge resources in different languages, even though users may not be able to understand linked and connected resources? Potentially, this could be an issue that frustrates users if they cannot “consume” the resources. However, the lack of awareness

of related resources in other languages could also have negative consequences.

Eighth, personalized portals, used so heavily in library settings can be used to facilitate the user's view, access and control of a broad range of multilingual resources. A future application at WAICENT can be determined the extent to which such capabilities are ultimately provided to a growing and more sophisticated user.

Ninth, multilingual ontologies are not a one-to-one issue. There is not always a unique term that corresponds to terms in other languages. Accordingly, as seen above, ontologies in different languages have different numbers of components. Similarly, keyword searches are not perfectly parallel in multilingual systems.

8. A model of multilingual KM systems

We present a model based on extending the approach of [19,20] to facilitate a categorization of multilingual approaches used by different firms and to facilitate our discussion of FAO and WAICENT. The model has three dimensions: “Conversion,” “Content” and “Connection.” In the case of multilingual systems, organizations can be somewhere along the spectrum of being completely passive in the “conversion” of knowledge resources to multilingual knowledge resources or actively converting knowledge resources to multiple languages. Similarly, knowledge resource “content” can be in a single language (“the official language is English”) or there can be content in multiple languages. If there are multiple languages, that content can be at the user interface or it can extend beyond the user interface to broad-based multilingual content. “Connections” between people and resources or resources and other resources, for example search or established links, can be based on a single language or reach out to multiple languages. These three dimensional considerations are summarized below in the diagram about multilingual knowledge management systems (Fig. 15).

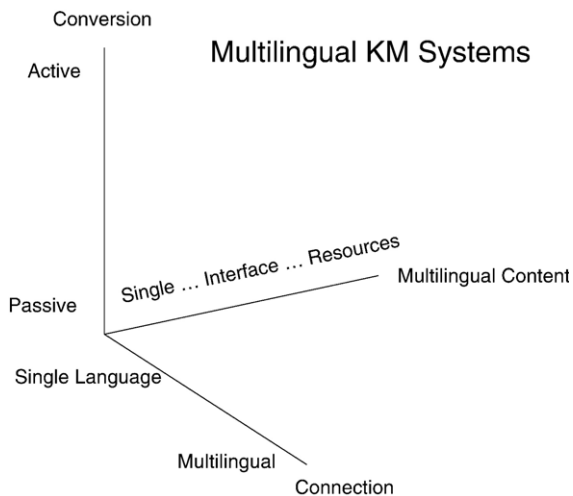


Fig. 15. Three Dimensions of Multilingual Knowledge Management Systems.

As an example, consider FAO and WAICENT in terms of this model. Clearly, FAO and WAICENT is an “active” converter. Knowledge resources are churned into multiple languages, depending on who needs the resources and developer capabilities. The multilingual content goes far beyond just a multilingual interface to include a broad range of multilingual resources as discussed above, such as ontologies and alerts. Connection, at FAO and WAICENT uses both single and multilingual approaches. In the case of search, as seen above, WAICENT provides multilingual connections. On the other hand, WAICENT Information Finder provides language specific connections between knowledge content. Based on this information we would categorize FAO WAICENT as being an active converter, multiple multilingual resources, and multilingual connections.

The model is quite robust and other systems also can be categorized according to this model. For example, Rozic-Hristovski et al. [25] apparently is multilingual along the “multilingual content” dimension, probably at the interface level. However, there is no reference to conversion or connection. As a result it would be on the multilingual content axis, somewhere in the middle. Sevinc [26] and others have stressed the need for medical research to be available in multiple languages. As a result, their notions of multilingual would likely be along the two dimensions of content and possibly conversion, with active conversion and available multilingual resources. As another example, Peters and Sheridan ([23], p.52) refer to “...accessing, querying and retrieving information from collections in any language ...” This would place their results along the two dimensions of multilingual resource content and multilingual connection.

The model might also be extended to other dimensions. For example, from our discussion here, we could consider multilingual search and other additional dimensions. For example, can search be focused on a single particular language or portfolio of languages or language independent search? As with content, this dimension becomes more important as we move to multilingual systems. However, the advantage of the current three dimensional approach is its visibility and simplicity, while focusing on key concepts.

9. Conclusion and contributions

This paper provides an analysis of a case of an organization that employs a multilingual knowledge management system. Initial analysis led to changes in the view of knowledge management: changing the need to “convert” knowledge resources and “connect” multilingual resources. As a result, this paper provides an analysis

of the key capabilities of a multilingual system actually used by millions of users. It is apparent that much planning and substantial resources have been put into play in order to have these systems work.

However, beyond the case study this paper provides insights into a number of other multilingual issues, including the following:

- User surveys can be developed using multiple languages in order to get feedback from a broad range of appropriate users.
- Multilingual can focus on the languages of on those countries affected (EMPRES), rather than just all the potential set of languages, in the case of limited resources.
- Multilingual ontologies or controlled vocabularies will not match up concepts one to one in different languages. Different numbers of concepts in each of ten different languages were found for the same controlled vocabulary (AGROVOC).
- In the analysis of the multilingual aspects of AGROVOC there appeared to be a “life cycle” for multilingual systems. The multilingual life cycle appears to be first to develop multiple language versions of the web page “around” the ontology. Second, the ontology itself is mapped into another language, while non-matching concepts are kept in the original language. Third, the documentation supporting the concepts is adopted last.
- Search by keyword can be by multiple sets of keywords, such as one per language (where the language could be chosen by the user) or using a single list of keywords that is itself multilingual (CARIS).
- Computer programs, such as web services, seem likely to have a name that is based on a single language (“getTermByLanguage”), however, using links and other devices it is possible to chose web services names that are in the context of the native language.
- The overall emerging architecture is XML-based, allowing substantial flexibility in the future.

Acknowledgements

I would like to acknowledge the many comments from the anonymous referees on two earlier versions of this paper. Further, I would like to acknowledge the many unique and insightful resources about FAO and WAICENT systems that made much of this analysis possible. The author apologizes for any errors and omissions that might have occurred, inevitable in a paper of this type. Finally, I would like to thank Christopher Yang for his work and patience with this project.

Appendix A. Selected Internet sites

- Nutch Home Page: <http://lucene.apache.org/nutch/>
 AGROVOC Ontology: http://www.fao.org/aims/ag_intro.htm
 UN Common Library: http://www.un.org.pk/library/unirr_waicent.html
 FAO Statistical Databases (FAOSTAT): <http://apps.fao.org/default.htm>
 EMPRES System: <http://www.fao.org/empres>
 Global Information and Early Warning System: <http://www.fao.org/giews/english/giewse.htm>
 Food Insecurity and Vulnerability Information and Mapping Systems (FIVIMS): <http://www.fivims.org>
 FAO Country Profiles and Mapping Information System: <http://www.fao.org/countryprofiles>
 International Information System for the Agricultural Sciences and Technology (AGRIS): <http://www.fao.org/agris/>
 FAO Document Repository: <http://www.fao.org/documents>
 FAO On-Line Catalogues: <http://www4.fao.org>
 Home-Page of FAO: <http://www.fao.org>
 WAICENT Access Statistics: <http://www.fao.org/wwwstats>
 Rome Declaration on World Food Security and World Food Summit Plan of Action, <http://www.fao.org/wfs/final/rd-e.htm>
 Home-Page of WAICENT: <http://www.fao.org/waicent>

References

- [1] Anonymous, The Agricultural Ontology Service, FAO, Rome, Italy, September 2001.
- [2] Anonymous, Info Finder Workshop Paves Way to Improved Searches on FAO site, http://www.fao.org/waicent/portal/detail_event.asp?lang=en&event_id=12643. June 06 2002.
- [3] M. Aiken, J. Martin, A. Shirani, S. Singleton, A group decision support system for multicultural and multilingual communication, *Decision Support Systems* 12 (2) (September 1994) 93–96.
- [4] K. Ciccone, “MyLibrary @ NCState: A library portal after five years,” *Journal of Library Administration* Volume 43, Numbers 1 / 2, pp. 19–35.
- [5] S. Cohen, J. Ferreira, A. Home, B. Kibbee, H. Mistlebauer, A. Smith, MyLibrary: personalized electronic services in the Cornell University Library, *D-Lib Magazine* 6 (4) (April 2000) <http://www.dlib.org/dlib/april00/mistlebauer/04mistlebauer.html>.
- [6] C. Goble, P. Crowther, D. Solomon, *A medical terminology server, Database and Expert Systems Applications*, vol. 856, Springer Verlag, 1994.
- [7] T. Gruber, A translational approach to portable ontologies, *Knowledge Acquisition* 5 (2) (1993) 199–220.
- [8] M. Jarrar, R. Verlinden, R. Meersman, Ontology-based customer complaint management, in: R. Meersmand, Z. Tari (Eds.), OTM

- Workshops 2003, Lecture Notes in Computer Science, vol. 2889, Springer - Verlag, 2003, pp. 594–606.
- [9] S. Katz, Multilingual Knowledge Management as a Strategy to Defeat Poverty and Hunger, www.schemas-forum.org/workshops/ws3/presentations/katz.ppt. May 10–11, 2001a Managing Schemas in a Multilingual Semantic Web, Budapest.
- [10] S. Katz, Content Management and Information Exchange, <http://www.bokis.is/iod2001/papers.html>. May 30 – June 1, 2001b paper presented at the 11th Nordic Conference on Information and Documentation, Reykjavik, Iceland.
- [11] S. Katz, J. Chelsom, A. Zisman, “Web Services in Action at the FAO of the United Nations,” XML 2002 Conference Proceedings, www.idealliance.org/papers/xml02/slides/chelsom/chelsom1.ppt.
- [12] J. Keizer, Issues, experiences and a project coherence and interoperability in agricultural information systems, International Conference on Dublin Core and Meta Data Applications, DC 2005, Madrid, September 2005, <http://dc2005.uc3m.es/program/presentations.asp>.
- [13] M. Last, A. Markov, A. Kandel, Multi-lingual detection of terrorist content on the web, in: H. Chen, et al., (Eds.), WISI, Lecture Notes in Computer Science, vol. 3917, Springer - Verlag, 2006, pp. 16–30.
- [14] B. Lauser, T. Wildeman, A. Poulos, F. Fisseha, J. Keizer, S. Katz, A comprehensive framework for building multilingual domain ontologies: creating a prototype biosecurity ontology, Proceedings of the International Conference on Dublin Core and Metadata for e-Communities, Firenze University Press, 2002, pp. 113–123.
- [15] B. Lauser, M. Sini, A. Liang, J. Keizer, S. Katz, AGROVOC and the OWL Web Ontology Language: the Agriculture Ontology Service — Concept Server OWL model, 5th Networked Knowledge Organizations Systems Workshop, Alicante Spain, September 2006, <http://www.ukoln.ac.uk/nkos/nkos2006/>.
- [16] R. Mayer, Navigation through terminological databases, Lecture Notes in Computer Science, vol. 898, Springer - Verlag, 1995, pp. 189–199.
- [17] E. Morgan, MyLibrary @ NCState, Proceedings of the Customized Information Delivery Workshop, SIGIR, Berkeley, CA, August 19 1999, pp. 12–18, <http://infomotions.com/musings/sigir-99/>.
- [18] D.E. O'Leary, Impediments in the use of explicit ontologies for KBS development, International Journal of Human Computer Studies 46 (1997) 327–337.
- [19] D.E. O'Leary, Knowledge management systems: converting and connecting, IEEE Intelligent Systems (May–June 1998) 30–33.
- [20] D.E. O'Leary, Using AI in knowledge management: knowledge bases and ontologies, IEEE Intelligent Systems (May – June 1998) 34–39.
- [21] C. Ohmann, H.P. Eich, E. Keim, “Multilingual decision support for the diagnosis of acute abdominal pain: an European concerted action (COPERNICUS 555),” in Artificial Intelligence in Medicine, Lecture Notes in Computer Science, Springer Verlag, # 1211, pp. 393–397.
- [22] I. Onyancha, J. Weinheimer, G. Salokhe, S. Katz, J. Keizer, Metadata exchange without pain, <http://dc2004.library.sh.cn/english/prog/ppp/james.ppt>.
- [23] C. Peters, P. Sheridan, Multilingual information access, in: M. Agosti, F. Crestani, G. Pasi (Eds.), ESSIR 2000, Lecture Notes in Computer Science 1980, Springer - Verlag, 2000, pp. 51–80.
- [24] J. Qin, Y. Zhou, E. Reid, C. Lai, H. Chen, Unraveling international terrorist groups exploitation of the web, in: H. Chen, et al., (Eds.), WISI, Lecture Notes in Computer Science, vol. 3917, Springer - Verlag, 2006, pp. 4–15.
- [25] A. Rozic-Hristovski, I. Humar, D. Hristovski, “Developing a multilingual, personalized medical library portal: use of MyLibrary in Slovenia,” Electronic Library and Information Systems, Volume 37, No. 3, pp. 146–157.
- [26] A. Sevinc, Multilingual approach to ‘Web of Science’, Journal of the National Medical Association 97 (1) (January 2005) 116–117.
- [27] J. Starr, Design considerations for multilingual web sites, Information Technologies and Libraries 24 (3) (September 2005) 107–116.
- [28] E. Verdu, M. Verdu, L. Regueras, J. De Castro, Intercultural and multilingual e-learning to bridge the digital divide, in: S. Shimojo, et al., (Eds.), HIS 2005, Lecture Notes in Computer Science 3597, Springer - Verlag, 2005, pp. 260–269.
- [29] G. Vouros, E. Eumeridou, P. Tselios, K. Kotis, Multilingual, ontology driven, content-based search and navigation of information items, Applied Artificial Intelligence 19 (2005) 691–719.
- [30] C. Yang, K. Li, An associate constraint network approach to extract multi-lingual information for crime analysis, Decision Support Systems 43 (4) (August 2007) 1348–1361.
- [31] Y. Zhou, H. Qin, CMedPort: An integrated approach to facilitating chinese medical information seeking, Decision Support Systems 42 (3) (December 2006) 1431–1448.



Daniel E. O'Leary is a Professor in the Marshall School of Business at the University of Southern California (USC). Dan has been at USC since 1985, when he got his Ph. D. from Case Western Reserve University. Recently, Dan has focused on Enterprise Resource Planning Systems and Knowledge Management. Professor O'Leary's book, Enterprise Resource Planning Systems: A Life Cycle Approach was published by Cambridge University Press.

Dan is a former editor of *IEEE Intelligent Systems*, and has been on the editorial boards of a number of journals, including *Decision Sciences*, *European Journal of Information Systems*, and *Expert Systems with Applications*. Professor O'Leary has published research in a number of different journals in different areas, including *Decision Science*, *European Journal of Operational Research*, *Management Science*, *Computer*, *IEEE Intelligent Systems*, *International Journal of Human Computer Studies*, *Communications of the ACM*, *Decision Support Systems* and *Journal of Management Information Systems*.