Agent-Based Up-to-date Data Management in National electronic Library for Communicable Disease

Patty Kostkova[1], Jane Mani-Saada[2], Gemma Madle[3], Julius Weinberg[4]

Institute of Health Sciences, The City University
London, Northampton Square, EC1V 0HB, UK
[1]patty@soi.city.ac.uk
[2] J.Mani-Saada@city.ac.uk
[3] G.C.Madle@city.ac.uk
[4]J.R.Weinberg@city.ac.uk

Abstract. Modern healthcare specialists are overwhelmed with medical information available on the Internet. However, it is difficult to find a particular piece of information when and where they actually need it. The National electronic Library for Health (NeLH) is addressing this issue by providing a single-entry portal to evidence-based medical information on the Internet enhanced with a quality tag assigned by professional experts in the field. In order to fully utilize the potential of an Internet-based library, the NeLH is distributed and consists of a number of Virtual Branch Libraries (VBLs), each dedicated to a particular disease or a medical area. Our team is responsible for the development of the communicable disease branch of the NeLH, called NeLCD (National electronic Library for Communicable Disease). VBLs are dynamically updated and their design reflects the needs of each particular user base. However, users accessing a single VBL may want to search the entire NeLH or should have the option of being able to search the entire NeLH. Therefore, support for a distributed search according to an adopted topology of VBL servers is essential. Intelligent Interface Agents are essential for the development and runtime of the library as they perform autonomously a number of tasks related to the search, assist humans in information publishing, the document review process and data exchange and retrieval. In this paper, we present an agent-based solution to assist in distributed search across the NeLH, and customization and personalization in the NeLCD.

1. Introduction

Modern healthcare specialists as well as general public are overwhelmed with medical information available on the Internet. However, they cannot find a piece of information when and where they need it [1]. Therefore, Department of Health (DoH) UK proposed a development of a National electronic Library for Health (NeLH) [1], a gateway to the evidence-based medical knowledge on the Internet. This paper focuses on agent-based development of the National electronic Library for Communicable Disease (NeLCD) [2], one of the virtual branch libraries (VBLs) forming the NeLH and the issues of distribution in the NeLH.

One of the crucial requirement of the NeLH is providing the *best available evidence* and *quality-tagging* the medical information for all user groups. By the term ”quality” we mean information giving the level of evidence, the source of the evidence, consistency, completeness, reliability and ”up-to-date-ness”. This will distinguish the NeLH from many other existing medical sites (such as Medline [3]) which typically offer comprehensive search for documents available on the Internet without giving the user an insight into the level of evidence and quality of the presented information [4].

One of the key applications of Intelligent Agents, a fast-growing area in software technology, is information retrieval and assistance in searching information in the Internet [5]. In addition, agents autonomous behaviour enables them to assist in various search-related processes on the users behalf, however, without human intervention. These features make them particularly exciting for health care applications, and in various health-related fields in general. However, there is no common agreement on ontology, nor agreed standards in health care (coding standards, data representation standards and common legal and ethical recommendations). For example, there is no common internationally accepted clinical coding scheme – currently, several coding systems are being used by different organizations: MESH, CTLV3/SNOMED and ICD10. This is not only a UK but an international issue.
The NeLCD, communicable disease branch library, investigates the application of Intelligent Agents in information retrieval, user customization and other aspect of the NeLH library, as being implemented in the UK.
This paper focuses on technical aspects of the distribution of the library into the Virtual Branch Libraries (VBLs), each dedicated to a specific area or a disease. There could conceivably be areas of overlap for example between the VBL for communicable disease and the VBLs for public health or primary care. The need for distributed search across the NeLH is essential in order to allow users to fully utilize the Internet-based resource. However, there is a tradeoff between flexibility and implementation simplicity.
This paper starts by giving a brief overview of the NeLH project (section 2), and then focuses on various aspects of the NeLH (section 3). Autonomous Intelligent Agents – supporting intelligent searching, enabling user profiling and managing the document review process, as being implemented in the NeLCD, are discussed in greater detail in section 4. Further, in section 5, our current work in progress is introduced demonstrating the usage of multi-agent systems (MAS) in information exchange within the scope of the entire NeLH project. Then, a discussion of various aspects is presented in section 6. Finally, in section 7 we discuss the project status, in section 8 the related work, and in section 9 we conclude.

2. The NeLH Background

The three major goals of the NeLH are [6]:

- To provide health care professionals and the public with knowledge and know-how to support health care related decisions

- To provide easy access to best current knowledge and know-how on the Internet
- To improve health and health care, clinical practice and patient choice.

The focus on clearly defined quality by stating a level of evidence will distinguish NeLH from other medical portals (Medline, PubMed, etc). In addition, the goal of NeLH is to provide the best available evidence, unlike many high-quality medical initiatives, such as Cochrane library, that provide the golden standard but omit many issues where the evidence is insufficient.

The main source of medical evidence are books, journals, and Internet-based sources. These include: Public Health Laboratory Service [7], Cochrane database [8], NHS Centre for Reviews and Dissemination, Effective Health Care Bulletins, British National Formulary, Centers for Disease Control, British Medical Journal, and others.

However, the quality, reliability and ”non-biasness” of provided information significantly vary. For example, studies of Mulrow [9], Oxman and Guyatt [10] have revealed how unreliable some editorials and review articles can be if they are not prepared systematically. In particular, although readers rely on journal review articles and editorials, the scientific evidence of these is inherently unreliable and biased towards a positive and optimistic view of the effectiveness of intervention [11].

The key approach supported by the NeLH framework is evidence-based healthcare that aims to clearly identify the level of evidence of a study or recommendation. As defined in [12]:

“evidence comes from a range of activities such as randomized controlled trials (RCTs), consensus statements, observation studies and surveys of patient views. Meta-analyses and large sample RCTs are considered to provide the highest quality evidence, with expert opinion and clinical experience at the lower end of the evidence hierarchy.”

Therefore, in order to overcome the problem of bias and unreliable medical evidence, and to ensure that the best current knowledge is delivered, NeLH aims to fully support evidence-based healthcare [13]. However, in order to achieve this goal autonomous agents need to be implemented to ensure that the available information is up-to-date, to manage and automate the documents review process and to respond to users specific needs.

The NeLH consists of virtual branch libraries (VBLs), each dedicated to a particular disease or a group of diseases. As has been said above, we are responsible for development of the NeLCD, the Communicable Disease branch of the NeLH, addressing issues of prevention, investigation and treatment of communicable disease.

3. Structure of the NeLH

NeLH is a single information gateway, a portal, to evidence-based information related to communicable disease with respect to all user groups – clinicians, GPs, public health professionals, environmental health officers, infectious control nurses, general

public and others. In this section, we will look at the general issues of the NeLH digital library and discuss internal data representation in the NeLCD.

3.1. Overview of the Library

The NeLH Knowledge Base is now a dominant commodity and knowledge distribution is so vital that the contribution made by computing, telecommunications and the World Wide Web are obviously of increasing importance. The development of knowledge management within healthcare is both necessary and possible [14]. It is based on existing medical evidence-based sites, PubMed, Cochrane Database, etc.
The core content of the NeLH can be likened to a central reference library and Virtual Branch Libraries to local branches, which serve the needs of the communities in which they based. So far, a number of prototype VBLs are being developed, such as Cancer, Child Health, Communicable Diseases, Diabetes, Diagnosis, Emergency Care, Health Informatics, Heart Disease, and others.
Our team is responsible for development of the NeLCD [35], the Communicable Disease branch of the NeLH, which addresses issues in prevention, investigation treatment and control of communicable disease. This is a single information gateway, a portal, to evidence-based information related to communicable disease with respect to all user groups – clinicians, GPs, public health professionals, environmental health officers, infectious control nurses, general public and others.
Figure 1 illustrates the NeLH framework. NeLCD stands for the National electronic Library for Communicable Disease, NeLC is the National electronic Library for Cancers, and the NeLPC is the National electronic Library for Primary Care.

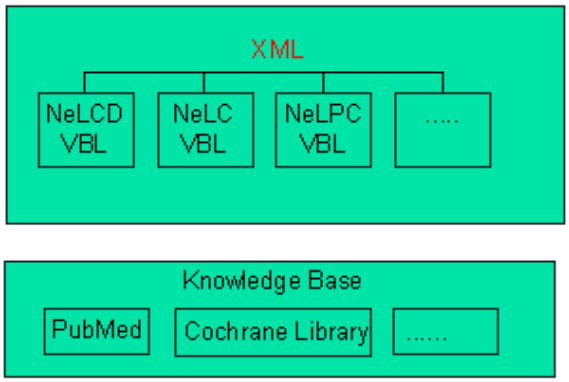

Fig. 1. The NeLH Structure

3.2. Knowledge Sources

The evidence available from NeLCD is obtained from high quality evidence-based directories, journals, and other databases. These include: PHLS [15], Cochrane

database [16], NHS Centre for Reviews and Dissemination, Effective Health Care Bulletins, British National Formulary, Centers for Disease Control, BMJ, and others.

3.3. Data Representation

In order to support a user-customizable search, documents in the library need to be precisely described. The Dublin Core Metadata initiative (http://www.purl.org/DC) defines a list of fields characterizing an electronic document for cataloging and search purposes. The NeLH adopted and extended this framework to better meet the requirements of quality and ”up-to-date-ness”. Issues related to metadata are covered in greater detail in the paper by Kamel Boulos describing a preparatory study of metadata, RDF and problem-knowledge coupling for the NeLH [17]. MESH medical ontology is used for indexing all documents in NeLH that is expressed in the field “Subject” of the Dublin Code metadata definition.
Every document in NeLH will be described by the following fields – called “electronic catalogue card”.

Table 1. Electronic calatogue card

Dublin Core Field	Description
Title	*Title of the document*
Creator	*Author(s) of the document*
Subject	*Keywords for indexing*
Publisher	*The publishing organization or the Internet site*
Date - publishing - posting - expiry	*Date of publication Date of NeLCD posting Date for review*
Type	*Publication type*
Format	*Software format*
Identifier	*Identifier (ISBN, URL)*
Source	*Document Bibliography*
Language	*Document Language*
Relation	*Reference to related documents*
Coverage	*Medical "category"*
Quality Tag	*Level of evidence*
Check List	*Answers to Check List Questions*
Description	*"Reviewers Assessment "*

	summarizing the document
ID	*Unique identifier*

3.4. The Appraisal Process

As each VBL serves specific group of healthcare professionals with particular information needs , there is a need for variations in the document appraisal process to meet the specific needs of each VBL's user base.
In this section, we will describe the appraisal process as it is set up within the NeLCD. The model of collaboration with professionals involved in the appraisal process may vary among VBLs, however, the key issues regarding the support for evidence and quality-tagging are agreed by all VBL teams.
As only pre-processing and quality-tagging of available information before incorporating it into NeLCD could ensure the required quality of the site, the core of NeLCD will consist only of appraised documents which have passed the NeLCD editorial procedure. This does not restrict the physical location of the actual documents – they could be local or accessed at their original source. Nor does it restrict the type of document available – the NeLCD attempts to present the "best available evidence", this could be a meta-analysis, or where there is little literature, a case report (but this is clearly indicated by the Level of Evidence, that is the Quality tag in the Dublin Core definition).
The Timestamp attached to every document in NeLCD will consist of three data items:

- Publication date (date of the document publication or "Access date" where no publication date is available)
- Posting date (date when document is added to the NeLCD database) and
- Expiry date (date when document is reviewed by the NeLCD team, usually 1 year after posting and annually there after).

As all information posted to NeLCD must be kept up-to-date, it is essential to review the core material regularly, even when no contradicting evidence has been found. For this reason every document has to go through a review process when its "Expiry date" (Table 1) has passed which is performed by the Expiry Agents. This is usually one year from the date of posting and then annually thereafter.
Each document considered the best available evidence on a particular subject is assigned a *quality tag*, by a member of a professional society or expert group.
The quality tag consists of:

- Level of Evidence (meta-analysis, Randomized Control Trial, Clinical Trial, Cohort Study, Case Control Study, Peer Leader Opinion, Personal Experience or Unspecified)
- Reviewers Assessment critically summarizing the paper, and
- Checklist, which answers brief questions about the methodological issues, level of evidence, potential biases and applicability of the results [18].

The resultant quality tag and a signature of the particular society are attached to the document and made available through the NeLCD.

4. Agents in the NeLCD

Agent architecture is described at two sections of this paper: firstly, agents supporting the information management within the NeLCD, and secondly, a set of agents responsible for distribution is discussed in the next section.
There are currently four basic agent concepts in the NeLCD. First two, Intelligent Search Agents and Pro-active Alert Agents, are involved in the search process and user profiling and customization. The later, Reactive Review Agents and Reactive Expiry Agents, are in change of various aspects of the library review process. The NeLCD is illustrated in Figure 2.

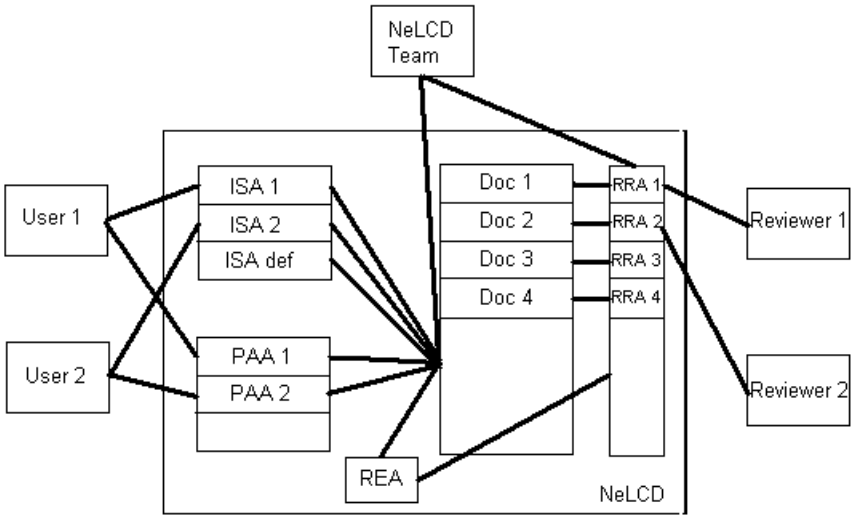

Fig. 2. Agents in NeLCD

4.1. Intelligent Search Agents

The primary goal of the library to provide a single portal for searching for an up-to-date medical information. Special type of Intelligent User Interface (IUI) agents [19], Intelligent Search Agents (ISA) are used to provide the search-related functionality.

Tasks

- Presenting users with a user-friendly adaptive interface to define his or her query for searching medical information in the library.
- The ISAs perform the searching functionality according to the given criteria.
- In addition to defining keywords for a query and their combination by logical operators, users can customize the search by restricting the resultant set to certain criteria, as indicated by the Dublin Core fields. For example, the agent can search for all documents regarding meningitis and children and display only Randomised Control Trials, or Case Studies.
- In addition, we are currently investigating a framework for adaptive ISA agents learning user profile – special interests, and preferences – to customize the interface and search capabilities to better meet user needs.

Number
There is an ISA for each registered user to allow the customization of the interface and the search,. In addition, as registration is not compulsory, there is a default ISA providing standard search and presenting default interface to non-registered users.

Interaction and Coordination
Each ISAs interacts with one user – that is, a customized ISA dedicated to a particular user provides the customized functionality, and a number of default ISAs serve non-registered users.
In addition, they also communicate with the Knowledge Sources to perform the search and return the results back to the user.
When a user logs in the site the dedicated ISA is activated and the user interface and search is customized accordingly. A new default ISA is created when a new user logs in.

4.2. Pro-active Alert Agents

Pro-active Alert agents (PAA) are performing tasks on users-behalf.

Tasks
- Users can instruct PAAs, to perform given tasks on their behalf autonomously or alert them accordingly. For example, the PAA is monitoring the library and informing users about new postings or comments related to specific issues.

Number
There is a PAA for each registered user to perform the alert functionality. Non-registered users have no PAA agents.

Interaction and Coordination
PAAs interact with users – one PAA per registered user. In addition, they communicate with the Knowledge Sources to check changes in the database interested to each particular user.

PAA act autonomously with a frequency given by the particular user – e.g., every day, once a week, etc. Users can set the PAA to alert about several issues, each with different frequency. PAA is in control of its activation according to the settings.

4.3. Reactive Review Agents

As a part of the appraisal process, experts from professional societies review each document by assigning it the quality tag. This process is managed by Reactive Review Agents (RRA).

Tasks

- RRAs are responsible for ensuring that each document is assigned a reviewer when its DC information is entered, and posting it to the library as soon as the review is delivered.
- In addition, the Review Agents remind the reviewers and the NeLCD team when the quality tag is not obtained within a given deadline.

Number

There is one RRA for each document in the database.

Interaction and Coordination

RRAs interact with reviewers and the NeLCD team (a special type of user) – a RRA sends off a request for a review to dedicated reviewer and wait for the reply. They send out another request and a notification to the NeLCD team if the review is no received within the given deadline. In addition, they interact with Reactive Expiry Agents (REA) which will be discussed bellow.

RRA have three states –

- "deactivated" (when the document is reviewed), is activated by entering a new document by the NeLCD team or by the REA
- "waiting – first request " (waiting for review when the first request has been sent out),
- "waiting – second request" (waiting for review when a second request and the information to the NeLCD team has been sent out). If the review is still not received, the NeLCD team needs to interfere to sort it out by choosing a different reviewer or finding an alternative solution.

4.4. Reactive Expiry Agents

Each document posted on the library is assigned an "expiry date" to ensure that all information is always up-to-date. This is the responsibility of the Reactive Expiry Agent (REA).

Tasks

- checking all documents for their expiry date
- the particular Reactive Review Agent is activated to inform the reviewer that the document needs to be reconsidered. It will take on the task and will remind the reviewer when the reply is not obtained within the given deadline, as is the case with new documents.

Number
The is only one REA in the database.

Interaction and Coordination
The REA collaborates with the Knowledge Sources to check expiry dates for all documents. In addition, it interacts with RRAs by activating them to take on the document review process. The REA runs autonomously once a week.

4.5. Example

There are some typical examples of the functionality of these agents as for the user perspective and document workflow perspective.

User perspective
Users typically set up their PAA by indicating the keywords defining their professional interests (based on MESH ontology), for example, meningitis in children & investigation, TB & treatment, Hepatitis A, etc. The PAA then searches the Knowledge Source for new entries related to these topics and sends an email to the user if a new document of his interest has been inserted.
In addition, users have their ISAs set up to search for information in the library. For example, the ISA is looking for documents on "journal papers" dealing with "prevention" of "HIV" as this is the user specialty. Obviously, the user setting could be overridden for any new search request. Then the ISA searches the Knowledge Source and presents the user the relevant documents. More advanced personalization, allowing users to customize the MESH-based keyword tree, is being investigated.

Document Perspective
Documents described by DC have phases they have to go through in the NeLCD that define the document workflow. Firstly, if a published document is considered the key evidence, it is included in the NeLCD and a new RRA is created. It sends an email to the appropriate professional society for request for the appraisal (the Reviewers Assessment). It is in the state "waiting – first request" until the Reviewers Assessment is submitted (by an online questionnaire) and then the agent go to the state "deactivated". Alternatively, it can request the RA once more if it was not received within a month – goes to the state "waiting – second request" which can lead into "deactivation" of the agent if the RA is received, Alternatively, an email to the NeLCD team is sent informing that the review is still pending, if there is no response from the processionals within another month.

Another core functionality performed by the REA, running once a week, is to ensure that all documents are reviewed again according to their "expiry date". The expired document's RRA is activated to submit the document for a review which follows the same process as in the case of a new document.

5. Distributed Search across the NeLH

The previous section of this paper have focused on interface and reactive agents related to medical information retrieval from the NeLCD. However, the NeLH contains a number of dedicated libraries (VBLs), as discussed at the beginning of this paper. The need for common ontology and data exchange specification as well as a framework for technical collaboration among the libraries is obvious.

The above section discussed technical aspects of the search for medical information within the NeLCD. Each VBL team is responsible for implementing the search facility within their virtual branch library, while accomplishing the requirements on data representation using the agreed Dublin Core-based electronic catalogue card. These are exchanged in XML format.

Medical ontology of data exchange and technical aspects of the interoperability among VBLs and the actual topology of VBLs are investigated in this section.

5.1. NeLH Topology

The NeLH architecture is being designed not as a simple client-server, but a set of VBL servers communicating with users who wants to search all available information within the NeLH. Therefore, the first technical issue to enable a distributed search is the topology of the digital library.

We are investigating a *star topology*, that is, a VBL processing a search will contact the NeLH server, which acts as a mediator, to obtain search results from all other VBLs. Another solution would be to implement a *Peer-to-Peer* communication where each VBL would separately contact every other VBL to receive all potential results.

We opted for the star solution for its implementation simplicity and easier consistency maintenance. The star topology of the NeLH is illustrated in Figure 3.

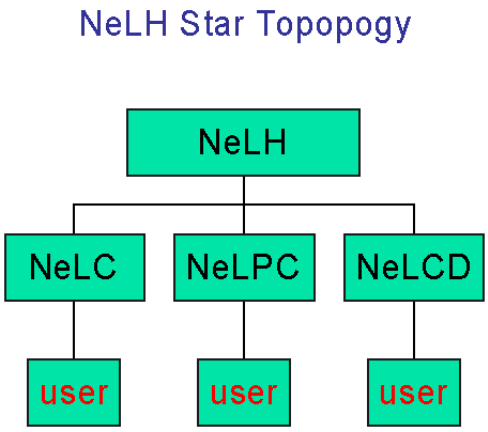

Fig. 3. NeLH Star Topology

5.2. Cross-search

As for the technical issues, NeLCD team is currently investigating the application of Distributed Cooperative Multi-Agent System (CMAS) [20,21] to the NeLH using a star topology. At each VBL, in addition to the Intelligent Agents discussed above, there are also Intelligent Search Agents for Distribution (ISA-D). In addition, there is the Central Intelligence Search Agent (ISA-C) at the top NeLH site [22]. These agents provide the additional functionality essential for the cross-search across the entire NeLH and are not needed within the NeLCD branch library – therefore, for reasons of clarity, ISA-C and ISA-D are introduced here and were not discussed in the previous section dealing only with a single VBL.

Intelligent Search Agents for Distribution (ISA-D)

ISA-Ds are present at each VBL and are responsible for ensuring the distributed search.

- They communicate with the particular user ISA to search the local VBL and, secondly, request results form remote VBLs via ISA Central (ISA-C) present in NeLH (discussed below).
- When the results are returned from the ISA-C, the ISA-D is also responsible for combining all results together (local and remote) and presenting them to the user.
- It is activated by a user request to receive all available documents in the NeLH, as the default is to search only the local VBL.

The Central Intelligent Search Agent (ISA-C)

The central NeLH site contains a ISA-C agent responsible for the following tasks:

- keeps up-to-date information of the location and availability of all VBLs (and references to the default ISAs).

- is responsible for combining the results coming from the VBLs ISA agents
- is responsible for filtering out duplications in search results before these are sent back to the user
- and sending the results back the end user, via his VBL ISA-D.
- They could be activated by any ISA-D in VBLs or directly by a user accessing the search facility at the top NeLH site.
- It also communicate with the Negotiation Agent present in NeLH, discussed in the next section.

The architecture is illustrated in Figure 4. Other agents are omitted for reasons of simplicity. The numbers next to communication lines illustrate steps in which the communication takes place.

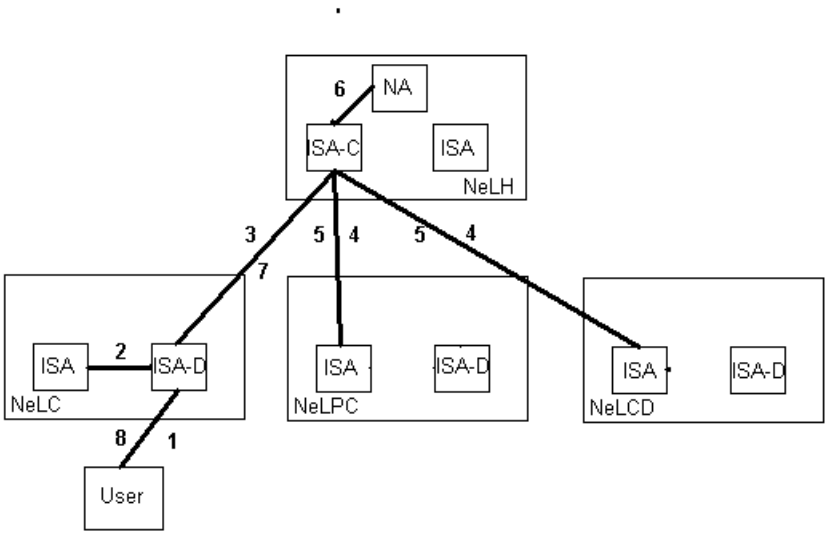

Fig. 4. Agents Performing a Distributed Search

Technically, an agent-based query producing an XML output is a robust and flexible solution for all technical platforms used by VBLs. The XML, defined by DTD/schemes based on the discussed DC, Table 1, could be rendered by the receiving VBL to meet the particular VBL requirements on look and feel, thought customisable ISAs, as discussed in [35]. In addition, server-side scripting should be adopted where possible to avoid browser-related problems. Therefore, the output XML sent to the receiving VBL for rendering is transformed into the resulting HTML which is then sent to the client.

5.3. Conflict-resolution

However, a single paper or medical document could come up in several VBLs and could be assigned different a quality tag in each VBL, as their appraisal procedure may vary.
There are two solutions to this problem. Firstly, there is a Negotiating Agent (NA), provided by the NeLH, which will autonomously step in and negotiate a common level of evidence. If agreement cannot be achieved without human intervention, the Negotiation Agent could request the check lists and an additional information from the VBLs involved in order to help the end user to make the final decision on validity of the evidence.
The second approach would be always presenting the user all search results, including those of contradicting quality in order to allow the professional to make the final decision. For these cases, providing the check list (which defines the criteria on which a certain level of evidence was assigned), is essential.
In addition, we are investigating the option of enabling users to customise their search so as they can define whether to authorize the NeLH to perform the negotiation, or whether they prefer to receive all contradicting results and decide themselves.

5.4. Example

For example, a user of the NeLCD is searching for "investigation documents on small pox of "Randomized Control Trial" level of evidence, however, is interested in evidence available in the entire NeLH, not just the NeLCD branch (where this would be performed by the ISA agent, as discussed above). In this case, the ISA-D is activated and calls the ISA to retrieve the relevant local documents. Then, it calls the ISA-C at the main NeLH server with the same request and waits for the result. In parallel, ISA-C, having the knowledge of the location and availability of all VBLs, sends a search requests to all ISAs in VBLs and waits for all responses. As for the ISAs in other VBLs, the search is performed in the same manner as if it was initiated locally. The results are sent back to the central ISA-C which can contact the Negotiation Agent to deal with possible inconsistencies in the review of the same document, however, the specification of this agent is currently being investigated. Then, all results are returned to the originating ISA-C that is responsible for combining the results and presenting them to the user. The numbers at the Figure 5 represent the order of these operations.

6. Discussion

The distributed star topology solution was investigated by the NeLCD team. Here we discuss other possible approaches to the library design.

6.1. A Centralized Library

A centralized library would be the best value for money. However, the users vary widely across the libraries (VBLs), and the libraries vary widely in scope and types of content – it is important to maintain the ability of each VBL to respond to the strategic needs of their user base. Therefore, an independence of VBLs in terms of implementing their database structure and search engine, as well as defining the look and feel of the VBL, the user interface and navigation strategy are essential. Also, having a direct control we have more flexibility to adapt and respond to our user base.

6.2. A Peer-to-Peer Topology

As has been discussed above, an alternative solution to the adopted star topology would be a Peer-to-Peer communication. The simplicity of maintaining consistency by the central NeLH server is the main reason for our approach. Therefore, there is just single point of information on VBL location, access, current availability, etc.
Also, all VBLs have to be contacted each time a cross-search is performed to receive all results. As the data are being frequently updated, it cannot be known in advance whether a certain VBL can provide information on a particular subject.
On the other hand, the drawback of this approach is that the NeLH server being as a single point of information is also a single point of failure. Therefore, in case of the NeLH server being unavailable, the cross-search, as discussed above, would not work.

6.3. Search Definitions

We believe that all user queries could be expressed as combinations of logical operators on columns (or full-text search), and therefore, the ODBS type of solutions are not needed. In addition, a number of the libraries have adopted the Lotus Domino database as their implementation platform, therefore, there is no need to provide cross-SQL queries as Lotus database is not relational.

6.4. Sharing User Profiles

Also, user profile could be exchanged among the VBLs to ensure the same functionality regardless of the actual VBL a particular user has logged in. The user could define his preferences regarding resolving conflicts, as discussed above, as well as user interface preferences. This area is also currently being investigated.

7. Project Status

Currently, we have finished our prototype phase and are developing the core library. The distributed search has been investigated by a number of VBL teams and currently we are in the process of finding a mutually convenient technical solution which will be implemented. As for the cross-search, currently, many VBLs are providing only a search of their local database, some support an explicit call to a server script to perform search in a particular remote VBL, such as VBLs which adopted standardized technical solution called the NeLH Toolkit [23]. However, a coordinated cross-search across all VBLs is still in the design and development phase.

7.1. Collaboration with Professional Societies

The major professional societies involved in the practice of communicable disease have indicated their support to the project and an initial model of collaboration have been set up. It is chaired by the Federation of Infection Societies, and secretarial support is provided by the Public Health Laboratory Service. The technical research and development of the project is provided by The City University [24].

7.2. Ethical Issues

Ethical aspects of any Internet-based healthcare project need to be addressed. In the case of NeLCD, no private patient data are gathered or processed. User profiles collected by Information Agents are kept and used with the owner consent. Personal data, such as names and e-mail numbers are kept by agents on local servers and not shared with agents in other VBLs. In addition, copyright aspects are not an issue either as the library provides links to documents which are already in the public domain [25].

7.3. Implementation Issues

The prototype of the NeLCD library (http://www.nelcd.co.uk) has been built using CGI scripts to implement the basic agent functionality. Also, many other VBLs have adopted proprietary solutions. The top NeLH site could be found at http://www.nelh.nhs.uk.
Currently we are porting the system to Lotus Domino R5 platform, which was chosen as the common environment for all VBLs in the NeLH. Lotus Domino built-in agents support is used for implementing the agent functionality discussed in this paper. The Lotus Domino agents provide the reactive and proactive functionality and are a substantial part of the development environment. Technically, agents are being written in Java Script and act as state machines being triggered by an event (user input – ISA) or a time definition (document expired - REA).

The prototype of the NeLCD library was built around the 10 high-priority areas, identified in a national prioritization exercise [26], and later populated with a core evidence-based documents with relevance to the infections diseases, syndromes and presentations. Documents in NeLCD are represented in XML (relevant DTD is defined for validation purposes). Common DTD across NeLH will allow extensive document exchange and cross-VBL search performed by Information Agents. The separation of the content from style allows flexible manipulation with data, easy modifications of the display format, as well as object-based data representation suitable for data search and document exchange. The documents available at the prototype site are in the editorial process now and the first Reviewers Assessments have been received.

Also, as some VBL teams are not fortunate in having technical expertise available, a toolkit platform has been developed [27]. This resource is essential for some teams, however, as has been said above, as a great number of libraries adopted a proprietary solution. Therefore, the interoperability issue is crucial to ensure data exchange and distributed search among all VBLs.

As has been said above, the physical distribution of VBLs and full control over technical development of each library is important to ensure flexibility and to respond to specific needs of each VBL users. However, in addition to a communication overhead, there are indications that this solution is not long-term viable for funding reasons. Therefore, it is a matter of discussion whether this project will realize the investigated distributed approach or whether less flexible solution would be the case.

7.4. Future Work

Currently, we are finishing of the MAS architecture described in this paper. In our future research, we will investigate the negotiation agents in greater detail, look at user profiling and more complex customization of ISAs and PAAs. Also, we will be looking at the personalization of the MESH-based ontology through ISA profiling allowing users to select keywords according to their specialty by pruning the MESH tree. Finally, we will investigate automatic download of new documents from PubMed based on data mining Intelligent Agents regularly reviewing new entries at PubMed.

8. Related Work

Autonomous gents providing various functionality in health care applications have been an interesting area of research in recent years in academia and industry. For example, agent project by Honeywell is investigating applications of autonomous agents in elderly patients nursing [28], agentcities-funded (www.agentcities.net) healthcare project is looking at implementing MAS for negotiating patients visits to specialists according to his or her condition and physical location [29]. Agents-assisted recommendation for screening of cancer patients and other projects were

investigated by Cancer Research Fund, UK [30]. However, the NeLCD seems to be the only project using agent technology in medical digital library.
As for the distributed communicate aspect of the library, there are a number related digital libraries providing a collection of cross-searchable documents in the Internet. The Z39.50 [31] standard specifies an abstract information system with a rich set of facilities for searching, retrieving records, browsing term lists, etc. At the server side, this abstract system is mapped onto the interface of whatever specific database management system is being used. The client application is unaware of the implementation details of the software hiding behind the network interface, and it can access any type of database through the same, well-defined network protocol. On the client side, the abstract information system is mapped back onto an interface which can be tailored to the unique requirements of each user. This provides a well founded universal solution, conceptually similar to the one in NeLH, however, the NeLH is a proprietary database which does not aim to provide universality. The same is the case for the general SDLIP communication protocol [32].
A similar approach, looking at a tree hierarchical topology for communicating agents was investigated by Kostkova as the MAGNET Architecture [33].
Also, commercial Web publishing products, such as developed by Interwoven [34], do not provide the additional autonomous functionality required by the NeLH document quality review and appraisal process, as discussed in [35].
Finally, a related non-agent knowledge representation project from NLM: the Unified Medical Language System (UMLS) develops and distributes multi-purpose, electronic "Knowledge Sources" and associated lexical programs to enhance systems focused on patient data, digital libraries, Web and bibliographic retrieval, natural language processing, and decision support. UMLS [36] includes a list of vocabularies in the UMLS Metathesaurus.

9. Conclusion

NeLCD, a virtual branch library of the NeLH, is developing an information gateway, a digital library, providing the best available evidence-based knowledge, enhanced with medical quality tags, to a wide spectrum of users: clinical experts, public health, general practitioners and general public. In this paper we have introduced this project and discussed the application of Intelligent Agents in information retrieval, user profiling, and the assistance in the documents review process in the NeLCD virtual branch library. In addition, we have introduced and discussed the star topology for NeLH enabling a distributed search across all libraries in the NeLH, based on MAS.

References

[1] Gray, J.A.M., S de Lusignan. *National electronic Library for Health (NeLH)*. BMJ, **319** , 1476-1479, (1999).

[2] P. Kostkova, J. Mani-Saada, J. R. Weinberg. *"Do you read the literature properly, and if you do how can others benefit from it?".* PHLS Annual Scientific Conference 2002, Warwick, UK, (2002).

[3] Medline. http://www.medline.com

[4] M.N. Kamel Boulos., A.V. Roudsari,, C. Gordon, J.A.M. Gray. *The Use of Quality Benchmarking in Assessing Web Resources for the Dermatology VBL of the NeLH,* Journal of Medical Internet Research 2001; 3(1): e5, (2000).

[5] N. J. Davies, R. Weeks and M. C. Revett. *Information Agents for the World Wide Web*. BT Technical Journal, **14: 4**, (1996).

[6] J.R. Weinberg, J. Mani-Saada, K. Smith. *The National electronic Library for Communicable Disease (NeLCD).* Poster. FIS Conference, UK, (2000).

[7] Public Health Laboratory Service. http://www.phls.co.uk

[8] Cochrane Database. http:// www.cochrane.co.uk

[9] C.W. Mulrow. *The medical review article: state of the science.* Ann. Intern. Med., **106**, pp 485-8, (1987).

[10] A.D. Oxman, G. H. Guyatt. *Guidelines for reading literature reviews*. Can. Med. Assoc. Journal, **138**, pp 697-703.

[11] NeLH Team. *The Knowledge and Know-how Platform*. Internal document (1999).

[12] Health Technology Board for Scotland. *Guidance for manufacturers on submission of evidence to Health Technology Assessments,* http://www.htbs.co.uk/docs/pdf/Consultation%20draft.pdf, (2002)

[13] Health Technology Board for Scotland. *Guidance for manufacturers on submission of evidence to Health Technology Assessments,* http://www.htbs.co.uk/docs/pdf/Consultation%20draft.pdf, (2002)

[14] The NeLH Architecture Paper. http://www.nhsia.nhs.uk/nelh/background/arch_index.asp

[15] Public Health Laboratory Service. http://www.phls.co.uk

[16] Cochrane Database. http:// www.cochrane.co.uk

[17] M.N. Kamel Boulos. *A Preparatory Study on the Dermatology Virtual Branch Library of the NeLH* (MSc Thesis in Medical Informatics), GKT School of Medicine, KCL, University of London, UK, (2000)

[18] http://www.staff.city.ac.uk/~jane/check.doc

[19] S. Green and F. Somers. *Software agents: a Review.* Intelligent Agents Group, Computer Science Department, Trinity College Dublin and Broadcom Eireann Research Ltd., http://www.cs.tcl.ie/research_groups/aig/iag/iag.html

[20]K. Sycara. *Multi-agent Infrastructure, Agent Discovery, Middle Agents for Web Services and Interoperation*. M. Luck et. Al. (Eds.): ACAI 2001, LNAI 2086, 17-49, Springer-Verlag Berlin Heidelberg (2001).

[21] M. Genesereth and S. Ketchpel: *Software Agents*, Communications of the ACM, **37**, No 7, pp. 48-53, (1994)

[22] P. Kostkova, J. Mani-Saada, J.R. Weinberg. *Distributed Search in National electronic Library for Health*. In the Proceedings of the IEEE DCADL, ICPP 2002, Vancouver, Canada, August 2002.

[23] http://www.nelh.shef.ac.uk/nelh/front.nsf/LL?openform

[24] J. R. Weinberg, J. Mani-Saada, P. Kostkova. *Development of a National electronic Library for Communicable Disease*. IDSA Conference, Chicago, USA, October 2002.

[25] P. Kostkova, J. Mani-Saada, J.R. Weinberg. Ethical Issues in Health Care on the Internet: "Much ado about nothing?", In the Proceedings of the Nordic interdisciplinary conference and workshop at NTNU "Making Common Ground: Methodological and Ethical issues in doing Internet research", Trondheim, Norway, June 2002.

[26] P. Horby, A. Rushdy, C. Graham, M. O'Mahony, et al. *PLHS Overview of Communicable Diseases 1999*. http://www.staff.city.ac.uk/~jane/ovcd.pdf

[27] http://www.nelh.shef.ac.uk/nelh/front.nsf/LL?openform

[28] Karen Z Haigh, John Phelps, Christopher W Geib *An Open Agent Architecture for Assisting Elder Independence.* In the Proceedings of the AAMAS 2002, Bologna, Italy, July 2002.

[29]Antonio Moreno, David Isern. *A first step towards providing health-care agent-based services to mobile users.* In the Poster session of AAMAS 2002, Bologna, Italy, July 2002.

[30]J. Fox. Understanding intelligent agents: analysis and synthesis In the Proceedings of Agents applied in Health Care, 15th European Conference on Artificial Intelligence, ECAI-2002, Lyon, France, July 2002.

[31]http://www.biblio-tech.com/html/z39_50.html

[32]http://www-diglib.stanford.edu/~testbed/doc2/SDLIP/

[33] P. Kostkova, J. A. McCann. *Inter-federation Communication in the MAGNET Architecture*. In the Proceedings of the Grace Hopper Celebration of Women in Computing Conference 2000, Cape Cod, Massachusetts, USA, (2000).

[34]http://www.interwoven.com/products/content_management/key.html

[35] P. Kostkova, J. Mani-Saada, J. Weinberg. *Agent-based Up-to-date Data Management in National electronic Library for Communicable Disease*. In the Proceedings of Agents applied in Health Care, 15th European Conference on Artificial Intelligence, ECAI-2002, Lyon, France, July 2002.

[36] http://www.nlm.nih.gov/research/umls/umlsmain.html