Agents in Online Healthcare Digital Library Management of the National Resource for Infection Control

Steve D'Souza
City eHealth Research Centre
City University
London EC1V 0HB, U.K.
+44 207 0404136

s.dsouza@city.ac.uk

Patty Kostkova
City eHealth Research Centre
City University
London EC1V 0HB, U.K.
+44 207 0404084

patty@soi.city.ac.uk

ABSTRACT

In this paper, we describe a specification for the use of agents within the National Resource for Infection Control (NRIC - www.nric.org.uk) in the UK.

The National Resource for Infection Control is an Internet medical digital library that provides a single point of access to quality appraised, evidence based information within the field of infection control.

One of the most important aspects of running an online healthcare digital library is the improvement of the functionality provided to the users, in order to facilitate the transfer of information in as useful a manner as possible. Software agents can help solve some of the problems involved in this process in an efficient manner. The methods by which agents may be used in this scenario to aid the improvement of this digital library are detailed here.

Keywords

Agents; digital library; healthcare; knowledge management; information retrieval.

1. INTRODUCTION

Modern day computing is a continuing struggle towards greater efficiency and increased productivity. This is true for both the software and hardware sectors of computing. Improvements within the hardware sector are mainly driven by improved manufacturing methods and new materials, whilst in a similar manner those within the software sector are also driven by new technologies and concepts. One such concept is the use of software agents - these can fully utilize their potential on the Internet - a distributed ever-changing computer environment.

There are many definitions out there of what a software agent is; many of these differ widely [1, 2, 3, 4]. However, to sum up the most commonly agreed upon aspects, we can say that a software agent is an autonomous, self-contained program, which is capable of making decisions and taking actions to reach its goals, based on its perceived environment [5, 6, 7].

These days, the availability of the Internet has added a wealth of information online about virtually everything, including medicine. More and more health care information is being made available online, not only journal papers, articles and presentations, but guides, policies, and treatment information, as well as much, much more. With such an overwhelming amount of information online, it can be difficult for healthcare professionals to find the information they need quickly, as they need it [8]. Online digital libraries are one answer, as they can piece together this knowledge, cataloguing it into a searchable, user-friendly format, to aid efficient information retrieval.

Medical digital libraries are key resources for enabling professionals to keep up to date on new developments in their field. These can allow users to share resources and experience to their mutual benefit [9]. Integrating these within an online web environment is useful for maximising this sharing capacity, by increasing the distribution of the library as well as allowing greater ease of use.

Agents can play an important part in many aspects of this knowledge management, from automating mundane tasks, to the more complex matters of anticipating user needs and acting in a proactive manner in relating the information to the user, rather than simply reacting to the user. They have already been proposed for use in tackling other problems in the healthcare sector [10, 11, 12].

This paper aims to demonstrate some of the ways in which software agents can improve the quality, efficiency and effectiveness of the National Resource for Infection Control making the whole experience more useful and more educational for healthcare professionals. The agent systems described here are in the development phase, and are currently being implemented on the Domino 6.51 platform.

2. HEALTHCARE DIGITAL LIBRARY
2.1 NRIC Overview

Funded by the Department of Health UK, the National Resource for Infection Control (NRIC) in the UK is a new online healthcare digital library being developed by and for infection control professionals, in response to the need for a single access point to existing resources for infection control, as indicated by reports from the National Audit Office [13].
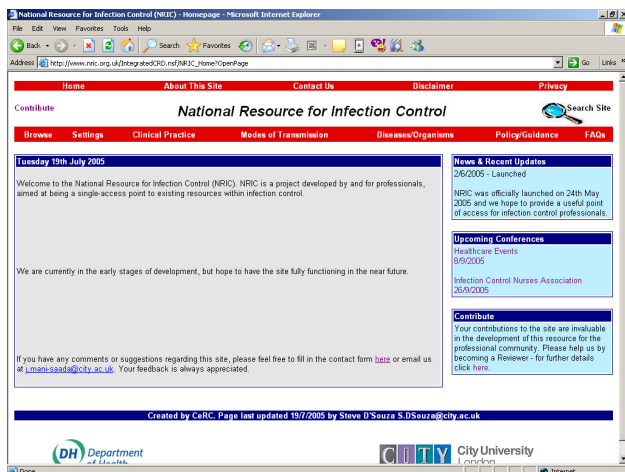
Figure 1. The National Resource for Infection Control website.

Policy and guidance documents are available on the website as well as templates to aid in writing local policies. Evidence based information is available and organised by settings, clinical practice tasks, modes of transmission and diseases / organisms. A search facility is also available for those who prefer not to browse. Each infection resource is quality appraised before being added to the website, using a set of questions assessing the levels of evidence, methodological soundness and the applicability of results, and many of the resources have a full review, written by infection control professionals. The implementation of online discussion of the resources allows further debate of a review, so as to ensure that all issues are addressed and any controversial reviews may be commented on.

2.2 Database Structure

In order to allow future compatibility for exporting or importing data, as well as to aid in user-customisable searching, the documents within the NRIC have a clearly defined format, based upon the Dublin Core Metadata initiative [14] but extended as appropriate to allow the expression of the quality review. Every document in the NRIC is described by the following fields:

Table 1. Document Fields within NRIC.

Field	Description
Title	Title of the document
Creator	Name of the author or organisation
Publisher	The publishing organisation or the internet site
Date	1. Date of publishing 2. Date of last review by publisher 3. Date for next review by NRIC
Type	Publication Type
Format	Physical or digital manifestation of the resource
Resource Identifier	Formal identification system (ISBN, URL)

Source	Reference to the resource from which the present resource is derived
Language	Language in which the document is written
Relation	Reference to related documents
Category	Medical category e.g. "Treatment" or "Prevention"
Level of Evidence	Evidence on which the document is based, e.g. "Clinical trial" or "Professional opinion"
Publication Type	Classification of the document type, e.g. "Factsheet", "Systematic Review"
Coverage	Regional application of this document, e.g. "National" or "International"
Abstract	Summary of the content of the resource
Access Rights	Information about who can access this resource or an indication of its security status
Quality Tag	Level of evidence
ID	Unique identifier
Review Status	Status of review process for this document, one of the following values: Unreviewed Being reviewed Reviewed Not applicable
Approval Status	Whether a document is approved for public viewing. Field values: Yes or No

New documents are automatically assigned "Unreviewed" as the default value for the Review Status field. As reviewers agree to review individual documents, this field changes for those documents to "Being reviewed" so as to flag this up and prevent others from unnecessarily reviewing the same document. Further reviews of a document may be added after an initial review, but as this initial review is likely to cover all the issues we are concerned with, it is preferable to get the next reviewer to spend their time reviewing "Unreviewed" documents, as this be more beneficial to the site as a whole. After a document review is submitted online and approved, the Review Status field is changed to "Reviewed".

The quality appraisal process includes the systematic process of allocating documents to the appropriate healthcare professionals to review, so that details of and a link to the main document may be displayed together with a full review. This allows users to access a brief "quality summary" of the document before they read it, allowing them to quickly decide whether or not this document is what they are looking for, saving them time in the whole typically very time-consuming information search process.

Discussion forums allow further comments and continuing debate on controversial reviews, so as to allow all opinions to be heard and reduce any bias, subjective views, omissions or errors.

3. AGENT USE IN NRIC

The two main areas within which we are investigating agent use in NRIC are:

1. Management of the review process
2. Personalisation

The review process as described earlier is a lengthy, time-consuming part of the NRIC document management if done manually (not that the review is time-consuming). It is an ongoing process, as new documents are continually being added to the library, and these need reviews. However, the healthcare professionals who review them are often busy and lack the time to devote to writing regular reviews, so these are often done on a now-and-then basis. Reminders are usually necessary to keep the reviews coming at a reasonable rate. Automating agent-driven data flow is needed for this process as it would free up more developer time for improving other aspects of the digital library.

Personalisation can be reactive or proactive, or a combination of the two. Users may customise information retrieval methods to suit their individual preferences, or the system may suggest its own customisations based on its outlook of the user. Clinician's usually lack the time to spend customising a system, and so a system which adapts itself to the clinician proactively rather than requiring the clinician to adapt it personally can be more useful. However different clinicians work in different ways, and so may prefer to see the data in different manners too, hence the most adaptive system would be one which allows both methods.

3.1 Agent Management of the Review Process

If we differentiate the reviewers into several distinct groups by specialty and area of expertise, then when a new document is added to the system, an agent allocates this document to one of the reviewers in the appropriate group (based on the topic of the document and the specialty information provided by reviewers at registration for the NRIC editorial network), and sends it (or a link to it) to this person for reviewing. This information is stored in a field in the document so that the system keeps a record of who is in the process of reviewing which document. Records are kept of how many documents have been reviewed by whom, and the number of documents each reviewer is currently in the process of doing. Then the next document of this type gets sent to the next reviewer in that group and so on. This way no individual receives huge numbers of "Review this" alerts at one go, as they are equally shared by the reviewers in that group. This process is performed by two agents working together as shown in Figure 2.
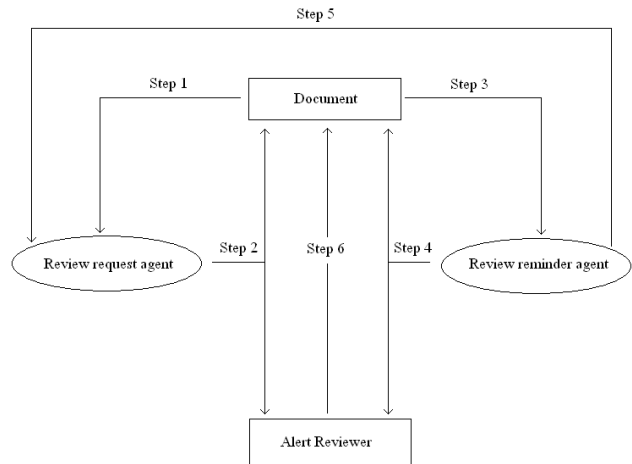

Figure 2. The Agent-based Review Management process.

Example Scenario:

A document is added to the database. Its default Review Status is "Unreviewed".

Step 1. The Review Request Agent selects documents which are "Unreviewed".

Step 2. It then sends a review request to the appropriate reviewer. It logs the reviewer name to a "Reviewer Request" field in the document, modifies the Review Status field to "Being Reviewed" and adds a timestamp to a "Review Request Date" field, also on the document.

Step 3. If 3 weeks have passed since the Review Request Date and the document is still "Being Reviewed", then this triggers the Review Reminder Agent.

Step 4. The Review Reminder Agent sends a reminder to the reviewer. It also logs this on the document in a "Reminder Sent" field which has values of "Yes" or "No".

Step 5. If 6 weeks have passed since the Review Request Date and the document is still "Being Reviewed", then this triggers the Review Request Agent. This agent then starts the process again, allocating it to the next reviewer in the appropriate group other than those logged in the Reviewer Request field.

Step 6. If a reviewer submits a review online, this automatically changes the Review Status field to "Reviewed", and ends the loop.

The review request gives the user the option of clicking a link to decline to review it, which then triggers the Review Request Agent to start the process again, or not doing anything to indicate acceptance. This way the reviewer has more flexibility and the system knows more quickly if the person isn't going to review it, in which case the agent passes it on to the next reviewer in line.

A third agent checks for documents that are about to expire within the next 4 weeks, and sends them back to be re-reviewed as necessary, using the same process as before. This involves

changing the Review Status to "Being reviewed" and triggering the Review Request Agent, with the request alert including a link to the existing review for modification if necessary.

3.2 Agents for Personalisation

Personalisation is an important use of agents that has also been of interest to others in the development of healthcare applications [15]. Personalisation of the site can be broken down into reactive and proactive personalisation. Reactive personalisation can include allowing the user to customise what they see by selecting certain topics to display information about and creating an individualised "home page" or set of pages. This would contain links to or embedded information on the topics or subsections of the site which the user chooses. This would of course require the user to login so that their preferences could be saved for future returns to the site.

Proactive personalisation is where agent technologies can play an important part. Information within the database needs to be presented in as efficient a manner as possible, so that access to key information is not overlooked by end users. Anticipation of what the user is looking for or may be interested in would enable a more productive information retrieval session. Agents can be used to analyse data trends of different groups of users, in order to find patterns in their information interests. A pilot study was performed on the National electronic Library for Infection in order to examine these patterns [16]; automated analysis would add further information on a continuous basis.

Agents can be used in this way to regulate personalised alerts. Alerts can include new documents, new reviewer's assessments, conferences, deadlines etc, and the agent can try to fulfil the user's wishes by matching the alert frequency with the selection of preferred intervals (i.e. daily alerts, weekly alerts, monthly digests, etc).

The personalisation alerts described here can be separated into three distinct types:

1. Alerts specific to profession (GPs, nurses, microbiologists, etc)

2. Alerts specific to specialty (e.g. Infection Control)

3. Alerts specific to individual choices of topics

The first of these is driven by agent ongoing data analysis of aggregate weblog data of each of those profession's as groups. The topics most common to each of these groups are selected for alerts so that the user can simply request the alerts typical for GP's for example.

The second type works in a similar manner, where the user can request alerts specific to and typical of a particular specialty.

The third type allows the user with the most time to customise their alerts to their individual preferences, choosing exactly which topics they want to receive alerts on, rather than the standard profession or specialty options.

Users in a group may have different interests, but also sufficiently similar interests for the development of group default preferences. These preferences can be predefined but may evolve over time as more use of the site provides greater feedback for analysis.

When a user registers for this service, they will need to specify both their profession and specialty, e.g. Nurse and Infection Control. As users browse the site whilst logged in, their browsing or searching patterns are recorded in the weblogs.
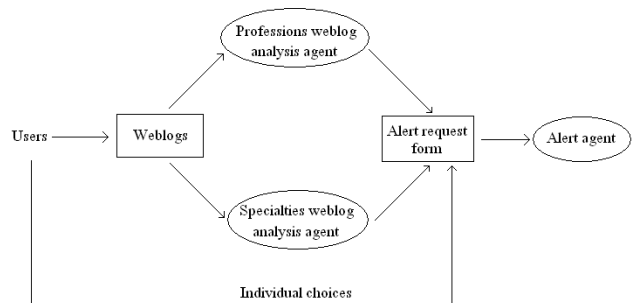

Figure 3. The Agent-based Personalisation alerts process.

Weblogs are stored for analysis of user behaviour in order to improve the site to better meet user needs. The information stored within the Lotus Domino weblogs includes IP address, username (if logged in), pages visited, and more. Weblog analysis provides useful information about the site such as the most commonly accessed pages, common search terms etc. Automated agent analysis of the weblogs can provide ongoing dynamic evaluation of the use of the site.

Having recorded the profession and specialty of each registered user as they first register to use the extended features of the site, two discrete user sets can be specified: Profession groups e.g. GP's, nurses, microbiologists; and Specialty groups e.g. Infection Control, Mycology, Tropical Medicine etc.

A Professions weblog analysis agent runs on a scheduled basis e.g. daily, to analyse the weblogs for users who have browsed the site whilst logged in. This agent ranks the topics viewed for each profession in the last 4 weeks, writing this information to a hidden field within the Alert Request Form. A second agent, the Specialties weblog analysis agent, runs in the same way to rank the topics viewed for each specialty in the last 4 weeks, and writes this to another hidden field within the Alert Request Form.

The Alert Request Form is used by users to specify what information they would like to receive in regularly mailed alerts. They can choose to accept the default setting of topics which are in both their profession and specialty rankings, topics in their profession rankings only, those in their specialty rankings only, or they can make an individualised choice from a complete list of all the MESH-based topics on the site. A mailing agent will then run

on a scheduled basis to email the personalised alerts to all those who have filled out the alert request form. This method is illustrated in Figure 3.

The data obtained by the agents through this profession / specialty weblog filtering can be examined further to also be useful for recommending information whilst the user is browsing the site. If, for example, a user is looking at a topic on the site whilst logged in, then the page will also show suggestions of other potentially interesting resources within the most popular topics of that professional group. Extra functionality can be added by the implementation of an underlying ontology to be a part of this recommendation, e.g. if a clinician is looking at information on TB, they will also be presented with a link to documents on HIV if an ontological relation between the two has been specified within the system.

All these methods of agent use can improve the usability of the site, and aid the efficiency of the whole knowledge management process of maintaining and using the online digital library.

4. Implementation experience

The National Resource for Infection Control (NRIC) is the test-bed for this agent use. The National electronic Library for Infection (NeLI) is the portal which hosts NRIC, and successful initial testing of this agent use within NRIC will lead to more widespread use of the systems across the NeLI and its other hosted projects.

In June 2005 NRIC had 184 unique users. As it was only launched at the end of May 2005, this number is expected to grow considerably as the site is promoted and gets greater exposure. In the same month the NeLI had 1475 unique users, whilst its other hosted projects had 6627. So the total number of users for NeLI and its projects in June was 8286.

4.1 Database Systems Used

The database system used to hold the core records that make up the NRIC digital library is IBM Lotus Domino 6.51. This is an integrated web application server, which can host web sites that a web browser, Lotus Notes clients and mobile clients may access, with the ability to serve pages stored in the file system or in a Lotus Domino database [17]. When a web browser requests a page from a Domino database, Domino translates the document record into HTML. Then the web server uses HTTP protocol to transfer the information to the web browser.

Lotus Domino's open, unified architecture has built in support for agent use. This includes web agents which run on the server, and can be event-triggered or scheduled. The agents must be signed by the owner and this digital signature in conjunction with server-wide settings determines the restrictions and rights of the agent.

Agents can be written for Lotus Domino in one of several languages: the built in Lotus formula language, Java or Lotusscript. For use on the web, agents in Lotusscript or Java can provide the greatest flexibility as they do not have the same access control restrictions as the formula language.

5. DISCUSSION & FUTURE WORK

Agent based methods have been increasingly generating interest within the healthcare sector [18, 19]. As autonomous intelligent entities which make independent decisions to reach their specified goals based on information received from within their environments, they can improve the functionality of a digital library, whilst at the same time reducing human maintenance.

Multi-agent systems offer a way of tackling distributed problems, but can also be used within a discrete environment to handle complex problems by dissecting them into separately handled issues and coordinating their efforts to solve them. This modular arrangement can be more reliable as no single point of failure exists that can crash the whole system.

With the huge amount of medical information available online, it is necessary to develop ways of accessing this information that are quick, easy and useful. Agents can facilitate the retrieval of information [20, 21] from multiple Internet sources using data mining techniques [22, 23] which is then analysed and filtered to suit individual user's preferences.

Work has also been done previously on weblog data mining techniques [24, 25, 26, 27] but most has been theoretical research and the application of these methods to healthcare digital libraries has been limited.

This framework described in this paper aims to apply this technology to a real world healthcare digital library which caters to the needs of a growing number of professionals within the UK. The main advantage of using agents within this library for the tasks described is that it allows a more flexible design, which can be easily tailored for use on the other NeLI projects as well.

There are other possibilities for agent use in the NRIC. One such possibility is the potential to use real-time agents for monitoring chat rooms and alerting a user if a topic matching the user's selected interests is currently being discussed. However running this sort of agent to monitor real-time chat rooms could be highly resource inefficient.

Another more practical use would be in using customisable agents to monitor the discussion forum threads for specific keywords or topics, as this would not have to be done in real-time and could be scheduled more efficiently.

6. CONCLUSIONS

In this paper we have described the specifications for intelligent agent based personalisation and review management within this National Resource for Infection Control (www.nric.org.uk).

We have demonstrated that there is potential for the use of software agents to add to the functionality of an online healthcare digital library such as NRIC, as well as improving efficiency and effectiveness in terms of both automating mundane tasks and in anticipating user needs through the analysis patterns within of aggregate data.

This technology can be helpful to both the developers and users, and similar methods may prove useful in other online information management applications.

7. REFERENCES

[1] Green S., Hurst L., Nangle B., Cunningham P.; "Software agents – a review"; https://www.cs.tcd.ie/research_groups/aig/iag/ (1997)

[2] Maes P.; "Intelligent software"; In Proceedings of Intelligent User Interfaces (1997)

[3] Nwana H.; "Software agents: An Overview"; Knowledge and Engineering Review, 11(3) (November 1996)

[4] Genesereth M. and Ketchpel S.; "Software agents" ; Communications of the ACM, 37, No. 7, pp 48-53 (1994)

[5] Nealon J., Moreno A.; "Agent-based applications in health care"; In "e-Health: Application of Computing Science in Medicine and HealthCare", Research on Computing Science Vol 5, Instituto Politecnico Nacional, pp 1-17 (2003)

[6] Jennings N.R.; "On agent-based software engineering"; Artificial Intelligence Vol 117, pp 277-296 (2000)

[7] Woolridge M.; "Agent-based software engineering"; IEE Proc. Software Engineering Vol 144(1), pp 26-37 (1997)

[8] Gray M., de Lustignan S.; "National electronic Library for Health (NeLH)"; *BMJ*, 319, pp 1476-1479, (1999)

[9] Marchionini G. and Maurer H.; "The roles of digital libraries in teaching and learning"; Communications of the ACM, 38(4), pp 67-75 (April 1995)

[10] Shankararaman V.; Proceedings of the workshop on agents in health care, at the 4th international conference on autonomous agents; Autonomous Agents 2000. Barcelona, (2000)

[11] Corts U., Fox F., Moreno A.; Proceedings of the workshop on agent applications in health care; 15th European conference on artificial intelligence, ECAI 2002, Lyon, (2002)

[12] Moreno A. and Garbay C.; Special issue of Artificial Intelligence in Medicine devoted to software agents in health care. (2003)

[13] Report by the Comptroller and Auditor General; The Management and Control of Hospital Acquired Infection in Acute NHS Trusts in England; UK National Audit Office (1999/2000)

[14] Dublin Core Metadata Initiative (DCMI); http://dublincore.org/

[15] Isern D., Sanchez D., Moreno A., Valls A.; Personalization of Medical Services.; In "e-Health: Application of Computing Science in Medicine and HealthCare", Research on Computing Science Vol 5, Instituto Politecnico Nacional (2003)

[16] Roy A.; "Results of the Online Questionnaire on the Evaluation of the National electronic Library for Infection"; http://www.nric.org.uk/Anjanareport3.nsf/Reportquest?Open Page

[17] IBM Lotus Domino; http://www.lotus.com/products/product4.nsf/wdocs/dominoh omepage; IBM Corporation

[18] Weiss G.; Multi-agent systems. A modern approach to Distributed Artificial Intelligence; M.I.T. Press (1999)

[19] Woolridge M.; An introduction to multiagent systems.; John Wiley, Chichester (2002)

[20] Klush M.; Information agent technology for the Internet: a survey.; *Data and Knowledge Engineering*, Vol 36(3), pp 337-372 (2001)

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

[22] Rebholz-Schuhmann D. and Kirsch H.; "Extraction of biomedical facts – a modular web server at the EBI (Whatizit)"; In the Proceedings of the Healthcare Digital Libraries Workshop (HDL 2004), (2004), Bath, U.K.

[23] Sousa P.A.C., Pimentao J.P., Garcao A.S.; "Intelligent agent technology: application to the health domain"; In the Proceedings of the Healthcare Digital Libraries Workshop (HDL 2003), (2003), Trondheim, Norway

[24] Ishikawa H., Ohta M., Yokoyama S., Nakayama J., Katayama K.; "On the effectiveness of web usage mining for page recommendation and restructuring".; Web, Web-Services and Database Systems, Revised Papers from the NODe 2002 Web- and Database-Related Workshops, Erfurt, Germany, (October 2002)

[25] Srivastava J., Cooley R., Deshpande M., Tan P.; "Web usage mining: Discovery and applications of usage patterns from web data"; SIGKDD Explorations, Vol. 1, No. 2, pp 12-23, (2000)

[26] Mobasher B., Cooley R., Srivastava J.; "Automatic personalization based on web usage mining", CACM, Vol 43, No. 8, pp 142-151, (2000)

[27] Spiliopoulou M., Faulstich L.C., Winkler K.; "A data miner analyzing the navigational behaviour of web users"; Proceedings of the Workshop on Machine Learning in User Modelling of ACAI99, (1999)