



Where Intelligent Technology  
Meets the Real World  
[www.pcai.com](http://www.pcai.com)

## Neural Networks

**Indexed vs. Unindexed  
Searching: From Security  
Classifications to Forensics**

**Neural Nets and Scientific  
Discovery: A Match Made in AI  
Heaven**

**The Visual Development of  
Rule-Based Systems**

**Protégé, Ontology and  
Knowledge Acquisition:  
Knowledge Representation,  
the Foundation of Intelligent  
Systems**

### **Also:**

**Business Applications,  
Data Mining,  
Data Modeling,  
Decision Support,  
Intelligent Search Tools,  
Intelligent Tools,  
Robotics,  
General  
Announcements,  
AI Conferences,  
Training,  
Fuzzy Logic,  
Fuzzy SQL,  
Intelligent Process Control,  
Neural Networks,  
Speech  
Recognition,  
Text Mining,  
Web Utilities**

## Table of Contents

### 18 Indexed vs. Unindexed Searching: From Security Classifications to Forensics

**Elizabeth Thede** discusses the differences between indexed and unindexed searching and when to use which technique.

### 24 Neural Nets and Scientific Discovery: A Match Made in AI Heaven

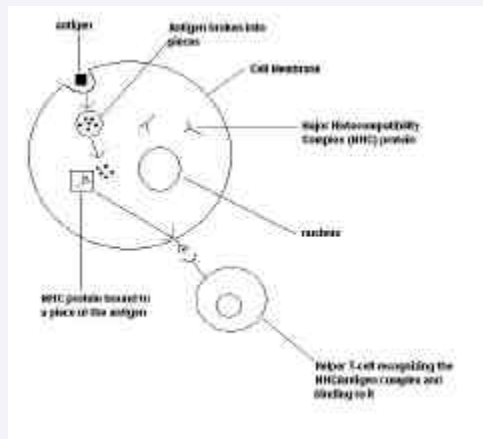
**Ilana Marks** presents some of the ways that neural network technologies are being used to expedite the scientific research process and provide valuable insights that might otherwise be overlooked.

### 29 The Visual Development of Rule-Based Systems

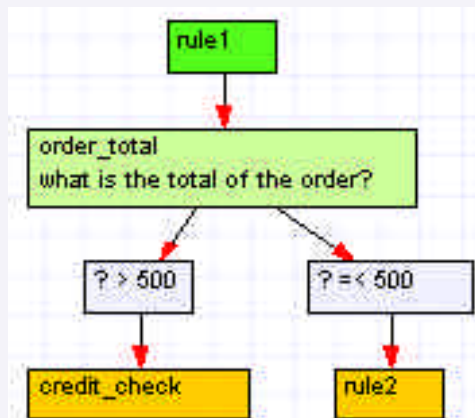
**Charles Langley and Clive Spenser** detail the creation of a rule-based system using visual representations of rules rather than traditional text-based representations.

### 37 Protégé, Ontology and Knowledge Acquisition: Knowledge Representation, the Foundation of Intelligent Systems

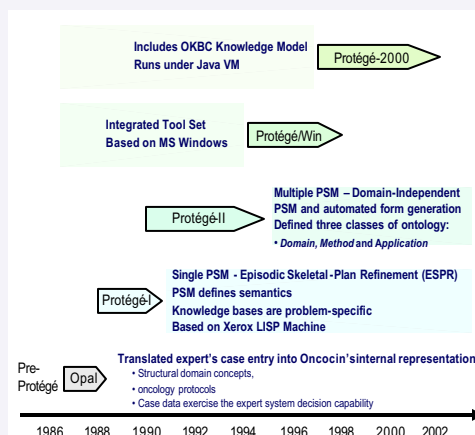
**Terry Hengl** discusses a knowledge-base application development tool called Protégé and how it uses ontologies to define terms and concepts in useful ways.



Page 24



Page 29



Page 37

## Columns

### 58 Book Zone

- \*Applying UML
- \*Text Mining: Predictive Methods for Analyzing Unstructured Information
- \*Spoken Dialogue Technology: Towards the Conversational User Interface
- \*Fuzzy Control of Queuing Systems
- \*Data Modeling Essentials, Third Edition
- \*Requirements Engineering
- \*Object-Oriented Construction Handbook: Developing Application-Oriented Software with the Tools & Materials Approach

**Ilana Marks**

### 62 Buyer's Guide

Fuzzy Logic, Fuzzy SQL, Intelligent Process Control, Neural Networks, Speech Recognition, Text Mining, Web Utilities

## Regular Features

### 7 Editorial

**8 Product Update** Business Applications, Data Mining and Modeling, Decision Support, Intelligent Search Tools, Intelligent Tools, Modeling, Robotics, Announcements, AI Conferences, Training

### 49 AI-Q Crossword Puzzle

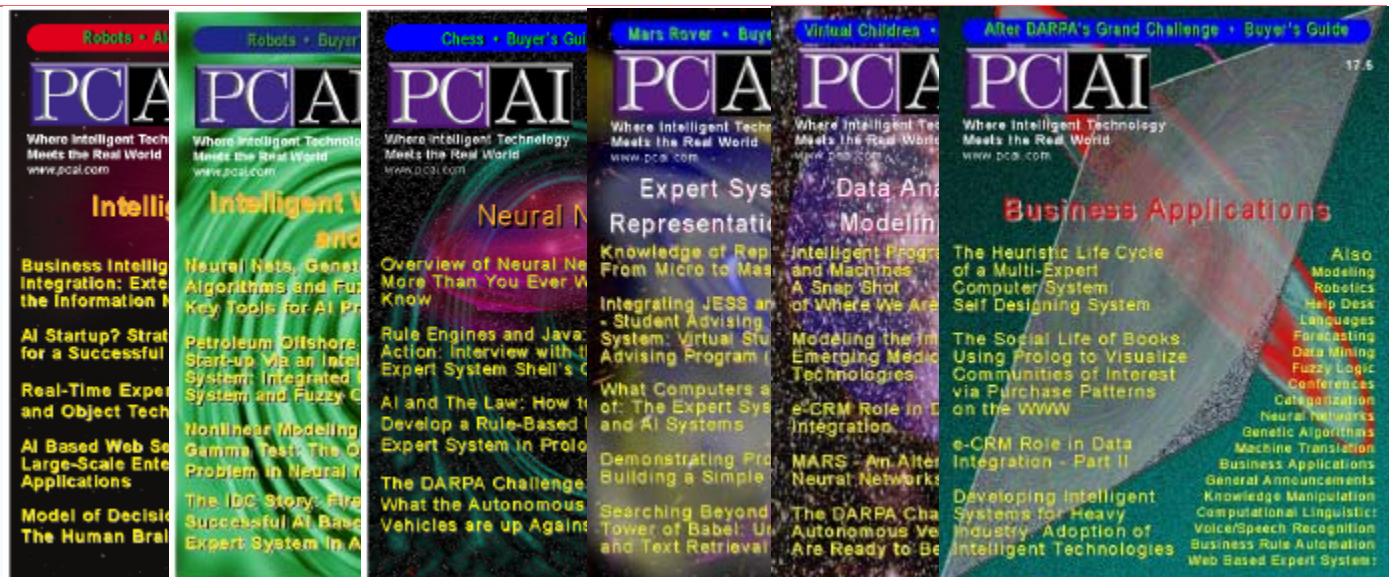
Neural Networks, Fuzzy Logic, and Speech Recognition

### 52 AI and the Net

Virtual Museum Education Assistant, Call Center Computers, Nose Mouse, Network of Robotic Telescopes, and more.

**Ilana Marks**

### 75 Classifieds and Recruiting



## **PC AI Volume 17 now on CD**

**6 Issues in both HTML and PDF  
Issues 17.1 - 17.6 (2003-2004)**

**Additional updates and bonus video not in  
the original issues.**

**All web issues reformatted with PC AI  
Navigation Bar**

**Higher Quality PDF's**

**Over 70 articles, columns and puzzles**

### **Sample articles include:**

*The Social Life of Books  
The DARPA Challenge  
Searching Beyond the Tower of Babel  
AI and the Law  
and much, much more.*

### **Topics included:**

*Data Mining, Robotics, Business Rules,  
Genetic Algorithms, Prolog, LISP, Business  
Forecasting, Case-Based Reasoning,  
Knowledge Management, Agents, Fuzzy  
Logic, Intelligent Tutoring, Expert Systems,  
Logic Programming, Modeling and  
Simulation, Searching, Pattern Matching,  
Natural Language Processing, etc.*

### **Buy PC AI Volume 17 on CD**

**For Paid Subscribers - \$14 each CD  
Non-Paid Subscribers - \$32 each CD**

**US Postage - \$3.00 for postage per CD**

**Foreign Postage - \$5.00 per CD)**

*(Quantity Discounts are available)*

### **Order online at**

**[www.pcai.com/store](http://www.pcai.com/store)**

**or**

**Contact PC AI directly**

**(602) 971-1869**



Also Available: Volume 16 CD  
Issues 16.1 - 16.6 (2002-2003)  
PLUS a special bonus issue -  
15.6 - AI to Combat Terrorism





Where Intelligent Technology  
Meets the Real World

**Publisher**

Terry Hengl

**Senior Editor**

Don Barker

**Webmaster**

Ilana Marks

**Columnists/Editors**

Elisa Hicks

Ilana Marks

Casper Goldberg

Paul E. Grayson

**Contributors**

Terry Hengl

Charles Langley

Ilana Marks

Clive Spenser

Elizabeth Thede

**Layout and Typography**

Michael Wiederhold

**Graphic Design and Illustration**

Laurens Watson

**Vice President of Marketing**

Robin Okun

**Editorial Assistant**

Lauren Dana

---

Please direct all editorial and  
advertising inquiries to:

PC AI

PO Box 30130

Phoenix, AZ 85046-0130

(602) 971-1869

[info@pcai.com](mailto:info@pcai.com)

[www.pcai.com](http://www.pcai.com)

Subscription rates are \$25 for one  
year (6 issues); \$40 for two years  
(12 issues)

Entire contents copyright ©2004 by  
Knowledge Technology, Inc., unless  
otherwise noted.

Authorization to photocopy items  
published in PC AI Magazine for  
internal or personal use, or the internal  
or personal use of specific clients is  
granted by Knowledge Technology,  
provided that the base fee of \$3.00 per  
copy, plus \$0.25 per page is paid  
directly to the Copyright Clearance  
Center.

\*PC AI (ISSN 0894-0711) is  
published bi-monthly by

Knowledge Technology, Inc.,  
PO Box 30130,  
Phoenix, AZ 85046-0130

# Editorial

## Learning to Communicate

One of the goals of artificial intelligence is to bridge the gap between the world of computers and the world of human beings. Enabling information to pass freely and easily between those two worlds can lead to many benefits. However simple a concept that may seem to be, in practice it is a difficult feat to accomplish. As we all know from yelling at our computers, they are not terribly good listeners. Also, when our computers try to communicate with us, the messages can be so complex as to be of little use to the average computer user. For example, if you have ever received an error message and clicked on the "show details" button, the resulting details look more like a Scrabble game that got out of hand than a helpful diagnostic of the problem. Therefore, a large hindrance to information exchange lies in the inability to effectively communicate. There is hope, though. With artificial intelligence technologies that attempt to emulate human processes, the wealth of information that can be generated by computers becomes more accessible to humans.

One such technology that is bridging the information gap is neural networks. These networks of interconnected processing units mimic the way that neurons in the human brain work. Important connections are emphasized while less relevant connections are downgraded. Neural networks can recognize patterns and predict possible outcomes just like humans can - but with the advantage of increased speed and capacity for information. As I discuss in my article "Neural Nets and Scientific Research: A Match Made in AI Heaven," this fact makes neural networks an invaluable tool in facilitating scientific research.

Another technology is intelligent searching tools. Searching through vast amounts of information can be daunting, especially if queries are taken too literally. However, computers, by nature, take everything literally. They use mathematical algorithms to evaluate problems, and are thus governed by the rigidity of mathematics. But there are ways to create more efficient and pertinent information searches, as Elizabeth Thede details in her article "Indexed vs. Unindexed Searching: From Security Classifications to Forensics." She presents the differences between indexed and unindexed searching and discusses how these enable organizations and individuals to search smarter and faster.

In their article "The Visual Development of Rule-Based Systems," Charles Langley and Clive Spenser discuss another problem with successful man/machine communication - how information is represented. Most people are more likely to understand a concept if it is presented visually, whether it be through diagrams, demonstrations, or gestures. In terms of rule-based systems, the information has almost always been presented in a text-based form. The authors contend that rule-based systems will be more effective if knowledge is presented in a visual form and they discuss the generation of such a system.

In Terry Hengl's article, "Protégé, Ontology and Knowledge Acquisition: Knowledge Representation, the Foundation of Intelligent Systems" he discusses a tool called Protégé which is designed to create customized knowledge-based applications. It works on the principle of ontologies which are definitions of concepts in terms of a language understandable to all parties involved. Ontologies also delineate relations between individual concepts so as to further define their meanings. This makes sure that everyone is "on the same page" and that knowledge is fully developed and useful.

As always, *PC AI*'s regular features are back. Test your knowledge of AI terms with the AI-Q crossword puzzle, learn about news in the artificial intelligence world with "AI and the Net," find a book or two to read in "The Bookzone," and discover new products with the "Product Update" and "Buyer's Guide." We hope you enjoy this issue of *PC AI* and learn something about the many intelligent ways that knowledge is conveyed in the information age.

**Ilana Marks**

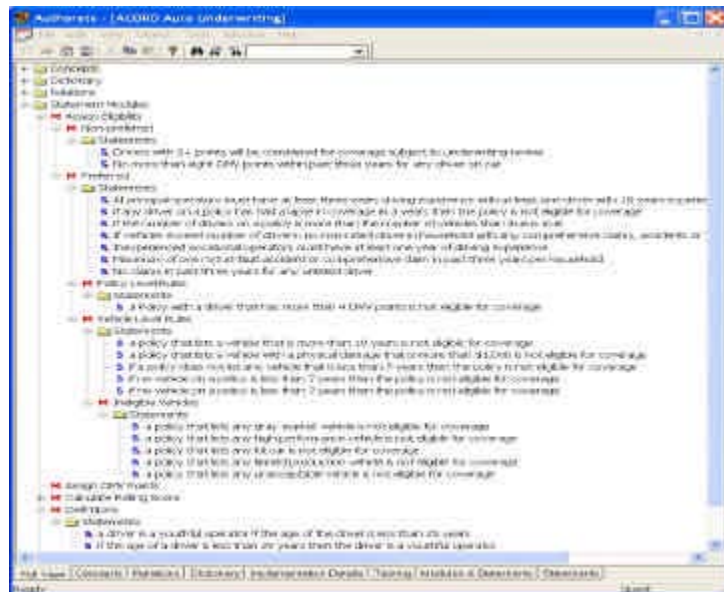
# Product Updates


## Business Applications

### HaleyAuthority 5.0 for .NET and Java Applications

Haley Systems, Inc. announces the release of HaleyAuthority 5.0, a business rules management tool for both .NET and Java environments. HaleyAuthority allows business and IT users to develop applications collaboratively while reducing the steps required to change rules when business conditions change. HaleyAuthority 5.0 offers natural language authoring and automatic code generation for .NET and Java applications, along with a patent-pending rules methodology. It offers multiple modes of rules authoring functionality which includes true English text, cascading menus, and table formats. HaleyAuthority 5.0 eliminates the need to translate business requirements from English into complex if-then logic and allows both business and IT users to manage business policies, procedures, regulations, constraints, etc. in an intuitive user interface.

Haley Systems, Inc.  
[info@haley.com](mailto:info@haley.com)  
[www.haley.com](http://www.haley.com)





# dtSearch®

## Instantly Search Gigabytes of Text Across a PC, Network, Intranet or Internet Site

### Publish Large Document Collections to the Web or to CD/DVD

- ◆ over two dozen indexed, unindexed, fielded & full-text search options
- ◆ highlights hits in HTML, XML, & PDF while displaying embedded links, formatting & images
- ◆ converts other file types (word processor, database, spreadsheet, email, ZIP, Unicode, etc.) to HTML for display with highlighted hits

**"The most powerful document search tool on the market"**  
*—Wired Magazine*

**"Intuitive and austere ... a superb search tool"** *—PC World*

**"Blindingly fast"** *—Computer Forensics: Incident Response Essentials*


**"A powerful arsenal of search tools"** *—The New York Times*

dtSearch "covers all data sources ... powerful Web-based engines" *—eWEEK*

**"Searches at blazing speeds"** *—Computer Reseller News Test Center*

In the past two years, over half of the Fortune 15 purchased dtSearch developer or network licenses.

**1-800-IT-FINDS** See [www.dtsearch.com](http://www.dtsearch.com) for:  
[sales@dtsearch.com](mailto:sales@dtsearch.com) ◆ hundreds of developer case studies & reviews  
 ◆ fully-functional evaluations



**The Smart Choice for Text Retrieval® since 1991**

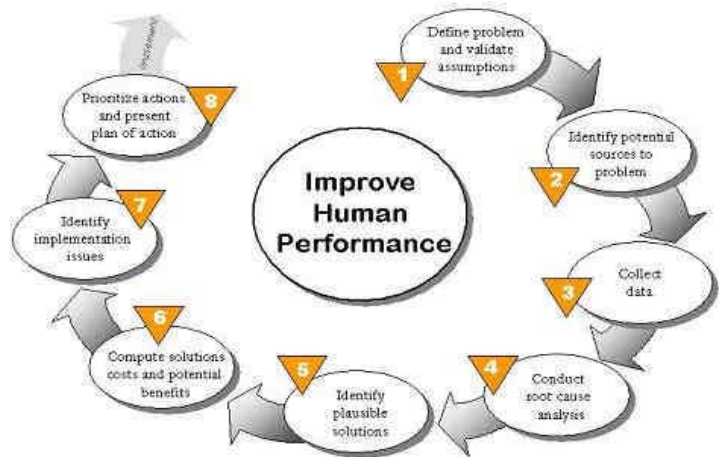


# Data Mining and Modeling

## Automated Data Mining Model Development and Validation Software

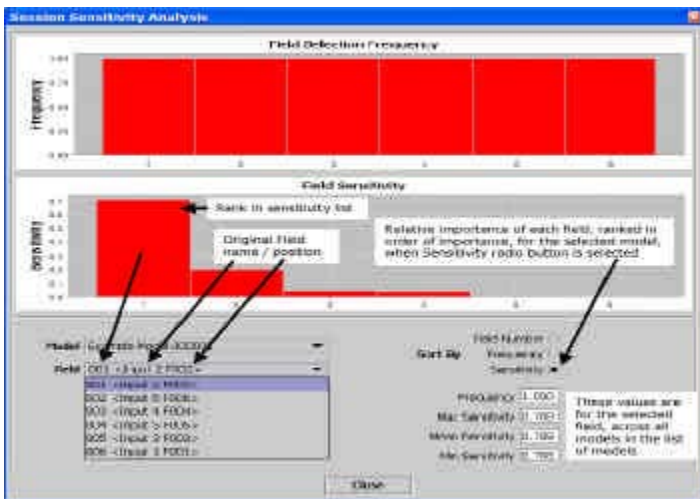
NeuralWare announces the release of the NeuralSight data mining model development/validation software. NeuralSight augments and enhances the NeuralWorks Predict® platform, which enables researchers and analysts to create neural network models for prediction, classification, and clustering applications. NeuralSight extends the features of Predict by providing a Graphical User Interface for handling large datasets and by allowing modelers to specify model performance requirements. NeuralSight selects and ranks the best neural networks from a set of networks built by Predict during an unattended model building session. NeuralSight is available now for Microsoft Windows® 2000 and Windows XP. Other operating systems will be supported in 2005.

NeuralWare  
[info@neuralware.com](mailto:info@neuralware.com)  
[www.neuralware.com](http://www.neuralware.com)



most cost-effective way to deliver the training. Module 4 is "Manage Training Budgets and Resources." This module determines how much money must be allocated to provide the most benefits. For more information on ADVISOR Enterprise, visit the BNH website.

BNH Expert Software  
[bahlis@bnhexpertsoft.com](mailto:bahlis@bnhexpertsoft.com)  
[www.bnhexpertsoft.com](http://www.bnhexpertsoft.com)



## Decision Support

### ADVISOR Enterprise 6.1

BNH Expert Software announces the release of ADVISOR Enterprise v. 6.1. ADVISOR Enterprise is a decision support tool that helps organizations manage training budgets and resources and identify methods of running programs more efficiently. ADVISOR contains four modules: Module 1 is "Align Training with Organizational Goals". This is a needs assessment tool for analyzing goals and identifying what is needed to achieve those goals. Module 2 is "Improve Human Performance." This is a decision support tool that analyzes deficiencies in performance and recommends training and solutions. Module 3 is "Select the Right Blend of Delivery Options." This module analyzes the effectiveness of training courses and then determines the

## INFERENCE WITHOUT THE ENGINE!!

**EZ-Xpert 3.0** eliminates the inference engine, so you get rule-based solutions with:  
**WARP SPEED** – test over 20,000 rules per second on a PC while using less memory.

**EZ, Fast Development** – the expert can create 100s of rules a day without assistance.

**Xpert Precision** - all applications meet 23 verification criteria, ensuring the highest accuracy.

**Free Distribution** - no royalties, *no inference engine to buy, learn, support, or implement.*

**You Write NO CODE!** – EZ-Xpert generates complete systems in *C++ and VisualBasic!*

## EZ-Xpert 3.0

*Let us port your expert systems into C++!*

AI Developers, Inc.

[info@ez-xpert.com](mailto:info@ez-xpert.com) (888) 327-9397

[www.ez-xpert.com](http://www.ez-xpert.com)

# Intelligent Search Tools

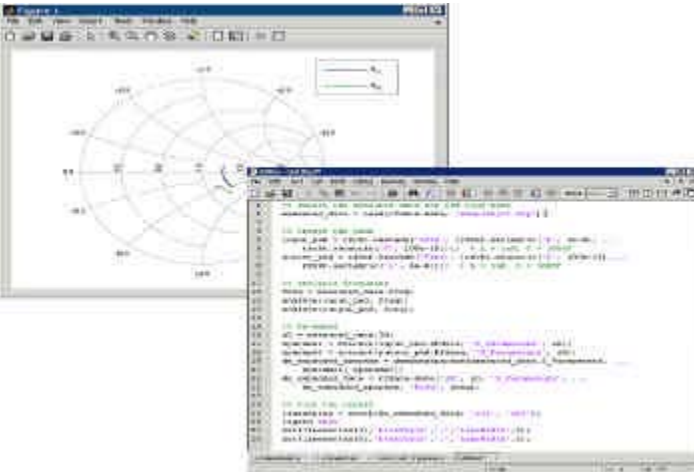
## New Version of DocuLex Desktop Search

DocuLex and dtSearch announce the release of Version 6.3 of DocuLex Desktop Search. The resulting product joins DocuLex's document capture technology with dtSearch's text searching technology. Desktop Search works with DocuLex's multiple OCR and other imaging products, including DocuLex's Professional Capture and Office Capture as well as Goby Capture. The integrated product works to index and search the full text of imaged documents. For example, after DocuLex OCRs a collection of documents

into PDF "image with hidden text" format, Desktop Search can then index and search that collection, highlighting hits on the scanned image. Desktop Search can also combine a search of imaged files with other PDF, HTML, XML, email (including attachments), word processor, database, spreadsheet, presentation and Unicode files. The product displays all retrieved files in a browser with highlighted hits, while keeping embedded XML, PDF and HTML formatting, links and images intact.

Doculex, Inc.  
[sales@doculex.com](mailto:sales@doculex.com)  
[www.doculex.com](http://www.doculex.com)

dtSearch Corp.  
[sales@dtsearch.com](mailto:sales@dtsearch.com)  
[www.dtsearch.com](http://www.dtsearch.com)



# Intelligent Tools

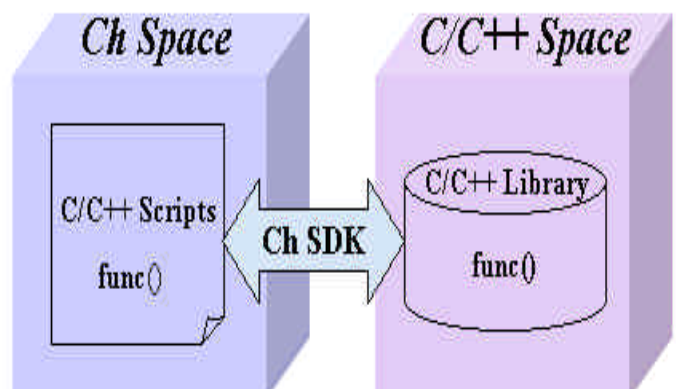
## MATLAB RF Toolbox 1.0.1



The MathWorks announces version 1.0.1 of their Radio Frequency (RF) Toolbox. The toolbox extends the MATLAB technical computing environment with functions and a graphical user interface for working with, analyzing, and visualizing the behavior of RF components. The toolbox allows the user to specify RF filters, transmission lines, amplifiers, and mixers by their network parameters and physical properties. The user can also read and write industry-standard file formats for network parameters. RF Toolbox functions are executed from the MATLAB command line or the RF Tool GUI, or can be called with individual MATLAB scripts and functions. The toolbox includes rectangular and polar plots and Smith® charts for visualizing data.

The MathWorks, Inc.  
[www.mathworks.com](http://www.mathworks.com)

## MPI-XF Support for AMD Opteron™ Processors

Engineered Intelligence (EI) announces the release of MPI-XF version 1.2, a high performance solution for the industry standard Message Passing Interface (MPI), and announced validation for




+


<p><b>With Eclipse:</b></p> <ul style="list-style-type: none"> <li>World class, multi-platform, extensible development environment</li> <li>Source code debugging for Prolog logicbases running in other programs and on remote (Web) servers</li> <li>Code outlines and project cross references</li> <li>Team development tools</li> </ul>	<p><b>With Amzi!:</b></p> <ul style="list-style-type: none"> <li>World class, multi-platform, extensible Logic Programming tools</li> <li>ISO Standard</li> <li>High performance for 24/7 server deployment</li> <li>Large logicbase support</li> <li>Integrates seamlessly with Java/JSP, .NET, C#, VB, C++, Delphi and more</li> </ul>
--	--

**You Can:**

- Automate the logical rules and relationships that run an organization
- Develop integrated logic base components
- Use professional development tools and methodologies
- Customize knowledge representation and reasoning engines for individual application needs
- Apply advanced ontology concepts and develop semantic web applications

**Download**  
[www.amzi.com](http://www.amzi.com)

# Indexed Vs. Unindexed Searching: From Security Classifications to Forensics

By Elizabeth Thede

Both indexed and unindexed searching have their place in the enterprise. Indexed text retrieval is typically more efficient for uses such as general information retrieval and security classification systems. But unindexed searching too has its place – in outgoing

email filtering, searching of live data sources like RSS news feeds, and sometimes in forensics. This article will attempt to explain which search technique to use when, and why.

## Overview: Indexed Text Retrieval

Indexing the inevitable millions of documents that any sizeable organization generates on shared file servers is the fastest way to facilitate data retrieval. An index will typically store each unique word in a document collection and its location within each document. Indexing also

### Sample Objects for Document Classification

In the dtSearch Engine, an "xfilter" can combine a full-text query with a filter for specific document attributes, such as file name, date, or size, or the presence in the document of a word or field. The field component can consist of a standard document attribute, or an attribute that dtSearch adds "on the fly" while indexing.

Search	Results
<i>(user request) and xfilter(name "abc*.html")</i>	This query would match any document that contains <i>(user request)</i> with a file name matching <i>abc*.html</i>
<i>(user request) and xfilter(word "projectxyz")</i>	This query would match any document that contains <i>(user request)</i> and that also contains the word <i>projectxyz</i>
<i>(user request) and (xfilter(word "Type::projectx") and xfilter(word "classification::high"))</i>	This final query adds two field restrictions to the <i>(user request)</i> : one for a named field called <i>type</i> with an entry of <i>projectx</i> , and the second for a named field called <i>classification</i> with an entry of <i>high</i> .

A dtSearch SearchFilter uses an in-memory object, consisting of a table of bit vectors, to achieve similar results to that of an xfilter.



works with non-document data, e.g. for forensics search purposes (see below).

After indexing, full-text search speed, even across millions of documents, is typically less than a second. While indexing a very large collection of documents for the first time may be time consuming, subsequent updates of the index are usually much faster. dtSearch, for example, simply checks the file modification dates of all indexed files, and only reindexes those files that have been added, deleted or changed since the last index update. (While the text retrieval terminology here relies on the dtSearch product line, the concepts in this article are generally applicable.)

In addition to enabling precision boolean searching,

an index can also store such information as word positions, enabling word or phrase proximity searching. An index can also hold information about word frequency and distribution, enabling computation of natural language relevancy rankings across a document collection. If the company name appears in two million documents, it would get a low relevancy ranking. If the latest marketing terminology appears in only four documents, it would get a much higher relevancy rank. In that way, PR could, for example, enter a whole paragraph of proposed text for a press release as a natural language search, and zoom right in on the most relevant documents.

But full-text searching, whether boolean, natural

language, or otherwise, is only part of the text retrieval answer. Suppose HR wants to limit its search to documents with an HR *executive* designation. This type of fielded data classification can result from fields or meta data inside a document, or from an overlaying document management-type application. With the latter, fielded data classification can rely on associated database entries, such as SQL or XML, or the addition of fields "on the fly" during the indexing process.

### Adding in Security Classifications

Now suppose the goal is to enable searching organization-wide, but to keep the wrong documents out of the wrong

(Text Continued on pg. 21)



4 out of 5 of Fortune Magazine's most profitable companies purchased dtSearch developer or multi-user licenses in the past two years.

# dtSearch®

## Instantly Search Gigabytes of Text Across a PC, Network, Intranet or Internet Site

**Publish Large Document Collections to the Web or to CD/DVD**

- ♦ over two dozen indexed, unindexed, fielded & full-text search options
- ♦ highlights hits in HTML, XML & PDF while displaying embedded links, formatting & images
- ♦ converts other file types (word processor, database, spreadsheet, email, ZIP, Unicode, etc.) to HTML for display with highlighted hits

**dtSearch Reviews...**

- ♦ "The most powerful document search tool on the market" — *Wired Magazine*
- ♦ "Intuitive and austere ... a superb search tool" — *PC World*
- ♦ "Blindingly fast" — *Computer Forensics: Incident Response Essentials*
- ♦ "A powerful arsenal of search tools" — *The New York Times*
- ♦ "Covers all data sources ... powerful Web-based engines" — *eWEEK*
- ♦ "Searches at blazing speeds" — *Computer Reseller News Test Center*

**dtSearch Desktop with Spider \$199**  
"Industrial-strength... superb" — PC Magazine

**dtSearch Text Retrieval ENGINE for Win & .NET for Linux**  
"Industrial-strength... superb" — PC Magazine

**dtSearch WEB with Spider from \$999**  
"Industrial-strength... superb" — PC Magazine

**dtSearch PUBLISH for CD/DVDs from \$2,500**  
"Industrial-strength... superb" — PC Magazine

**dtSearch NETWORK with Spider from \$800**  
"Industrial-strength... superb" — PC Magazine

**1-800-IT-FINDS**  
sales@dtsearch.com

See [www.dtsearch.com](http://www.dtsearch.com) for:  
♦ hundreds of developer case studies & reviews  
♦ fully-functional evaluations

**The Smart Choice for Text Retrieval® since 1991**

# Neural Nets and Scientific Discovery: A Match Made in AI Heaven

By Ilana Marks

## Introduction

The human brain is an invaluable and amazing organ. Electrical and chemical messages bouncing around in the brain lead to actions, speech, thoughts, and all other physical and mental properties that define the human condition. One of the most interesting abilities that the brain affords us is the ability to recognize things, people, scents, and other physical stimuli. When you think about the diversity even within a certain classification of object, it becomes more amazing. For instance, within the classification of tree, there is a wide variety of different types of trees, many of which look drastically different. A pine tree looks very different from an oak tree and yet we realize that they are both trees. We recognize general characteristics of a tree and then are able to expand our definition when other specimens are encountered. We also recognize different states in the life of a tree - for example, we recognize that leaves may change color or fall off. Despite such seasonal anomalies, a tree is clearly recognizable.

In addition to the ability to recognize things, the brain also allows us to predict outcomes and occurrences. When dark clouds appear in the sky, we can predict that rain is imminent (unless, of course,

the weather report says it will rain - in that case, the sky will suddenly become clear!) Sports fans try to predict the outcome of a game based on the previous performance of the teams. Given a set of circumstances, it is possible to deduce the most likely result. Typically, the larger the set of circumstances available to make a prediction on, the more likely it is to come up with the correct results.

Recognition and prediction are both highly involved in scientific research. Since science is based on other science it is important to recognize where a result is reminiscent of a previous discovery. Recognizing those connections allows the researcher to expand their understanding by applying information garnered from previous research. A cornerstone of good scientific research is the formulation of an appropriate hypothesis. A hypothesis is nothing more than a prediction about what will occur. Therefore, prediction is a backbone of research. Scientists are constantly striving to increase their knowledge base in order to make more accurate predictions about new experiments.

## Neural Nets and Research

Recognition and prediction may

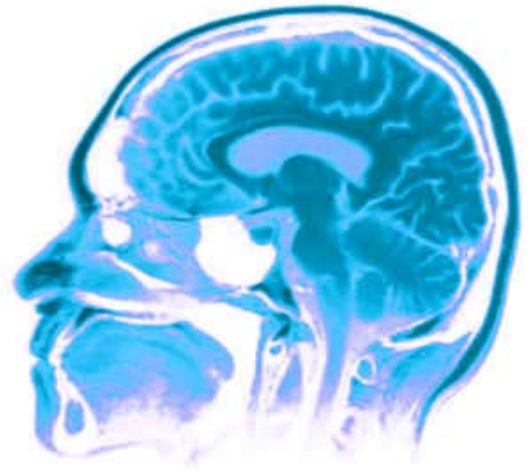


Figure 1: The human brain

be important to good research, however in this day and age, speed is almost more important. Universities demand that professors publish papers often and the public demands new drugs and technology. The time needed for scientists to learn all about a variety of different fields makes it prohibitive for them to make truly accurate predictions. In addition, a certain type of scientist may not recognize a certain occurrence because it may fit more in the knowledge base of another type of scientist. With the vast increase in scientific discovery over the last several decades, a mountain of knowledge is quickly being created



and its growth shows no signs of stopping. The human brain simply cannot deal with so much information. Therefore, it is necessary to employ artificial intelligence techniques in the process of scientific research. Artificial intelligence provides increased speed and "capacity" compared to the human brain, which proves incredibly useful to researchers.

Specifically, neural networks are a very useful tool. Neural networks mimic the behavior of the human brain, so they are often used in applications where systematic thought processes are important. Processing units are interconnected and communicate like biological neurons. Successful and relevant connections are emphasized whereas irrelevant connections are downgraded. This mimics the conditioned learning of biological organisms. This dynamic learning is vital in the research process since science does not remain static.

## Cancer Prediction

Cancer is a difficult disease to research. Cancer results from everyday

cells that, in effect, "lose their mind." In normal, healthy cells, the process of cell replication is highly regulated. Many "checkpoints" are in place to make sure that all steps are done properly. These "checkpoints" are usually molecules that are created or degraded which signal the cell to either hold in the stage where it is or to move on to the next stage of cell division. For example, one such molecule believed to be involved in several different cancers is called p53. This molecule is often absent or mutated in extracts from tumor cells, indicating that perhaps p53 is a molecule that signals the cell to stop dividing. If p53 is not produced, then cells will divide too rapidly and cancer can result.


Since no two cancer physiologies are exactly alike, it is notoriously difficult to make a prediction about whether the patient is likely to survive. Even with two cases of the same type of cancer, one person may survive for many years while another will die within months. A myriad of different physical factors contribute to the outcome. This is one reason why cancer is such a frightening

disease - no one can tell you whether your particular complement of genes will put you at an advantage or disadvantage.

Researchers at the National Cancer Institute (NCI) are working on creating a model that will help predict the prognosis of cancer patients. Particularly, they are working with a type of cancer called Neuroblastoma. Neuroblastoma is a childhood cancer. It usually begins with cells of the adrenal gland and spreads, creating tumors in the neck, chest, abdomen or pelvis. Using DNA microarray analysis (see PC AI Volume 18, Issue 2 - "Microarrays and Artificial Intelligence") the researchers studied the gene expression profiles of cancer patients. The microarrays consisted of about 25,000 genes and the analysis was repeated for 49 patients. In order to connect the microarray results with certain outcomes, the 49 patients were chosen because their outcomes were known. Some of the patients survived for more than three years without any cancer-related issues. Others died due to the disease. Feeding this information into an artificial neural network, the researchers found they


## New Exsys Selector RuleBook®

### Knowledge Automation Expert System Software Brings Interactive Product Selection to the Web.




Interactive  
"product expert"  
recommendations  
for each  
individual.

Hurry,  
only \$995.



- Fast and affordable - systems can be built in a few hours
- Logic is stated in the context of product selection problems
- Results provide individual customized reports
- Development environment uses simple notebook style interface
- Match look-and-feel of existing Web site
- Ideal for diagnostics and probabilistic identification
- Deploy as Java applet, or with optional Servlet Runtime
- Easily migrates to Exsys CORVID

FREE 30-Day Eval - [www.exsys.com](http://www.exsys.com)



EXSYS®

Capture Knowledge  
Deliver Answers™

# The Visual Development of Rule-Based Systems

By Charles Langley and Clive Spenser

## Introduction

In the late 1980's Knowledge Based Systems (KBS) were seen to be leading edge software technology. Developers thought that the simplest KBS paradigm, Expert Systems, perhaps combined with probabilistic and fuzzy logic extensions would soon revolutionise the way that software was used throughout business and other sectors of the economy.

KBS software was built on rules which encoded the knowledge of experts in any given domain. Computers would then use this encoded knowledge to make decisions on behalf of their human users.

It was not long however, before the bubble of hype surrounding these systems began to burst. Something was wrong, but what was it?

## The Knowledge Acquisition Bottleneck

Apart from the limited power of the computers available at the time, the major problem was the difficulty of acquiring implicit knowledge from the minds of experts and then representing it explicitly. This so-called Knowledge Acquisition Bottleneck was believed to be the limiting factor on building systems that could do complex, useful tasks.

By the end of the twentieth century however, university departments were working hard at this problem. Curiously it was often Psychology departments rather than Computer Science departments which had the most impact in this area.

In particular, Ethnography (by then seen as a core part of Cognitive Psychology) was being used to study behaviour in situ with the aim of identifying the cognitive processes underlying that behaviour. Just as Margaret Mead

(an early ethnographer) had lived amongst native tribes in Papua New Guinea in order to study their cognitive behaviour, so Psychology departments were sending researchers (often under cover) into workplace environments to discover how people approached problem-solving activities.

This work was, and continues to be, very successful. Knowledge acquisition is no longer the 'black art' it was deemed to be. Despite this, KBS has continued to be underused. Why might this be?

## A Knowledge Representation Bottleneck?

It is my contention that the problem was not primarily with how we obtained knowledge, but with how we represented it. I am not arguing that rules (or Bayesian networks and other knowledge representation methods) are inadequate to the

Order the new Volume 17  
CD at [www.pcai.com/store](http://www.pcai.com/store)



task, but rather that it is the way in which these rules and other representational formalisms are themselves represented that is the limiting factor.

At first a simple rule-base is relatively transparent, especially if properly documented. Certainly such systems were easier to comprehend than procedural code and were subsequently easier to update and amend. As such rule bases became larger and more complex however, a simple syntax error, perhaps only involving one word, could prevent them from operating correctly. The complexity of these rulesets also meant that it was difficult to get an overview of what was intended, thus impeding their maintenance and extension.

## A Picture is Worth a Thousand Words

The problem of rulebase comprehensibility, I would argue, is the fact that we have primarily represented knowledge using text based structures rather than visual ones. No matter how close to natural language a knowledge representation language is, you cannot see at a glance what a complex system is trying to do.

## Visual Rule Generation

Rule generation via a graphical interface is a hot topic right now, with offerings from a number of companies small and large. This is being driven in part by current interest in so-called 'business rules management' which is arguably a reawakening of the KBS paradigm we mentioned earlier.

London based Logic Programming Associates is an appropriate company to enter this market as it has been producing rule-based software since the mid 1980s. Its latest product, VisiRule, enables rule-based systems to be automatically generated from a flowchart drawn on the screen.

Consider the following business rule (Ross 2003):

Rule: An order must be credit-checked if any of the following is true:

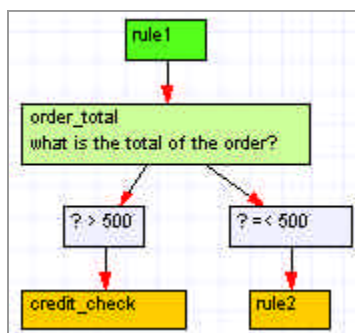


Figure 1

- \* The order total is more than \$500
- \* The outstanding balance of the customer's account plus the order amount is more than \$600
- \* The customer's account is not older than 30 days
- \* The customer's account is inactive
- \* The customer is out of state

Build more powerful,  
scalable applications  
with  
**Allegro CL**  
and **Lisp**

**Webify Seamlessly**

Deploy applications over the  
internet with AllegroServe and  
other web/app servers

**Easy Connectivity**

Integrate applications with a  
variety of Java, XML, database,  
CORBA and COM/OLE tools

Get to market *on time* and  
*on budget* with  
**Allegro CL and Lisp.**

Find out why so many  
leading companies & top  
researchers won't program  
with anything else.

**Free trial download!**

**www.franz.com**

AllegroServe is a trademark, & Allegro CL is a registered trademark of Franz Inc. Copyright 2002 Franz Inc. All rights reserved.

# Protégé, Ontology and Knowledge Acquisition: Knowledge Representation, The Foundation of Intelligent Systems

By Terry Hengl

## Introduction

Knowledge is a decisive competitive-advantage for today's corporations. Knowledge of schedules, raw materials, labor, manufacturing and distribution is essential to the supply chain while knowledge of customer interests and buying habits, latest technologies, budget constraints, marketing plans are crucial to product development. It offers a powerful tool for gaining market share and preserving a competitive edge, but it is costly to capture and control. Methodologies and technologies that assist in the acquisition, maintenance, and distribution of knowledge are essential to an organization's success. Today's society, and the world in general, have contributed to this growth in importance of knowledge management and knowledge based systems. Some examples include ([www.worldedreform.com/intercon2/f15a.pdf](http://www.worldedreform.com/intercon2/f15a.pdf)):

- \* Accelerating rate of change in every aspect of technology and society
- \* Staff migration and attrition (downsizing and reengineering)
- \* Geographic dispersion associated with globalization of markets
- \* Global integration of cultures, companies and markets
- \* Increase in networked organizations
- \* Increased level of education and

training of the population

- \* Growing knowledge-intensity of goods and services
- \* Revolution in information technology

It is not easy to efficiently and cost-effectively identify, acquire and maintain this knowledge. At a minimum, organizations must be able to:

- \* agree to an organization-wide vocabulary to ensure knowledge is consistently communicated and understood;
- \* identify, explicitly represent and model this knowledge;
- \* share and reuse this knowledge across independent applications and domains.

This article looks at these aspects of knowledge acquisition by examining Protégé (<http://Protégé.stanford.edu>), a free open-source Java tool with an extensible architecture for creating customized knowledge-based applications - based on Ontologies ([www.ontology.org](http://www.ontology.org)). It also reviews the concept of Ontologies (see side bar), and associated support methodologies, which establish the vocabulary and model the concepts along with their inter-relationships. This concept also includes processing of the associated attributes for a particular field of knowledge. By reviewing the evolution of Protégé, an ontology modeling and knowledge acquisition

environment (developed by Stanford Medical Informatics (<http://camis.stanford.edu>) at the Stanford University School of Medicine), an understanding of fundamental concepts that underpin knowledge acquisition emerges. This environment creates and modifies the ontologies and knowledge bases it generates to enable developers and domain experts to build knowledge-based systems.

Other associated technologies and standards mentioned in this article include the Open Knowledge Base Connectivity (OKBC), Generic Frame Protocol (GFP), Resource Description Framework (RDF), OWL and OWL-S.

## The Evolution of Protégé

Opal ([http://smi-web.stanford.edu/pubs/SMI\\_Abstracts/SMI-86-0137.html](http://smi-web.stanford.edu/pubs/SMI_Abstracts/SMI-86-0137.html)) was an expert system shell-based application developed as part of the medical domain Oncocin (<http://citeseer.ist.psu.edu/context/1419258/0>). Oncocin developed this knowledge acquisition and advice system for protocol-based cancer therapy. Opal enabled patient history entry by the

Order the new Volume 17  
CD at [www.pcai.com/store](http://www.pcai.com/store)



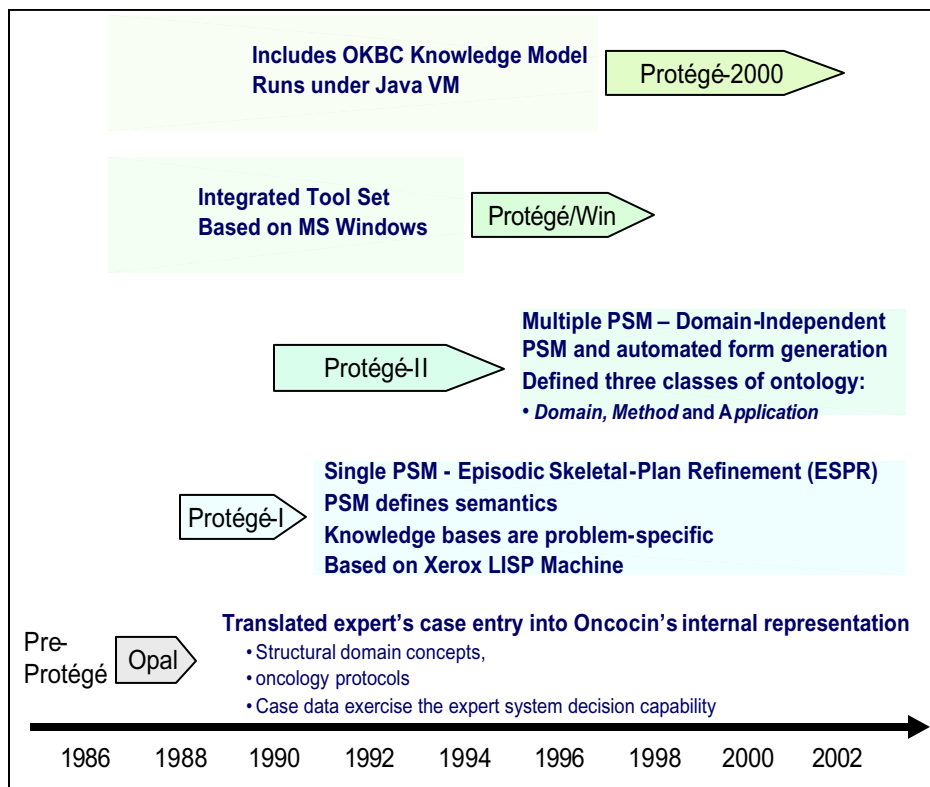


Figure 1: Evolution and enhancements of Protégé from 1986 to present.

physician or nurse (the domain experts), resulting in a suggested treatment or test. The knowledge base, a collection of if-then rules and other data, captured the clinical protocols. Opal translated the expert's input, via graphical forms, into an internal representation specifically tailored for Oncocin. This project identified three different levels of knowledge for this particular information:

1. Structural domain concepts used by the knowledge engineer to create the Opal knowledge-acquisition application;
2. The domain expert (oncologist) knowledge - oncology protocols;
3. Case data entered by the user to exercise the expert system decision capability.

Since Opal was inference engine based, it enabled reuse by knowledge engineers to create different knowledge bases - ultimately creating domain specific expert systems. The knowledge engineer was responsible for the knowledge acquisition; a tedious and time-intensive task. Unfortunately, the concept of a knowledge engineer separated the domain experts from the domain knowledge bases and this separation introduced a potentially large source of incorrect knowledge.

In 1987, Mark Musen built an

application for knowledge-based systems with a goal of building knowledge-acquisition application as part of Oncocin. Based on these three different levels of knowledge, he believed that knowledge acquisition occurs in phases with knowledge obtained in one phase defining the structural knowledge for subsequent phases. Musen's goal was to reduce the work engineers did to construct knowledge bases during knowledge-acquisition. He noticed that knowledge obtained during a specific phase influences the knowledge related application required for later stages.

From this early concept, Protégé evolved through four

phases, resulting in a rich development environment available for both research and knowledge-management. (see [http://smi-web.stanford.edu/pubs/SMI\\_Abstracts/SMI-2002-0943.html](http://smi-web.stanford.edu/pubs/SMI_Abstracts/SMI-2002-0943.html) for more information on the evolution of Protégé.

## Protégé-I - Knowledge Bases

This early phase, where Protégé-I simplified the knowledge acquisition process for building medical expert systems, learned from the earlier OPAL-based system. The intent of the tool was to simplify the knowledge acquisition process for the knowledge engineer, already overloaded with complex tasks to perform - a key issue with early expert system development. One goal of Protégé was to provide an application that created Knowledge Acquisition tools (KA) from a formally defined collection of concepts. This enabled the domain expert to create the knowledge base, eliminating the time consuming process of the knowledge engineer learning the domain. Early assumptions for Protégé 1 included:

**Inference without THE ENGINE**

**EZ-Xpert 3.0**

**Features:**

- Warp Speed - Test over 20,000 rules per second
- Fast Development - Easy to learn, create hundreds of rules a day
- No need to write code
- 23 Verification Criteria insure accuracy

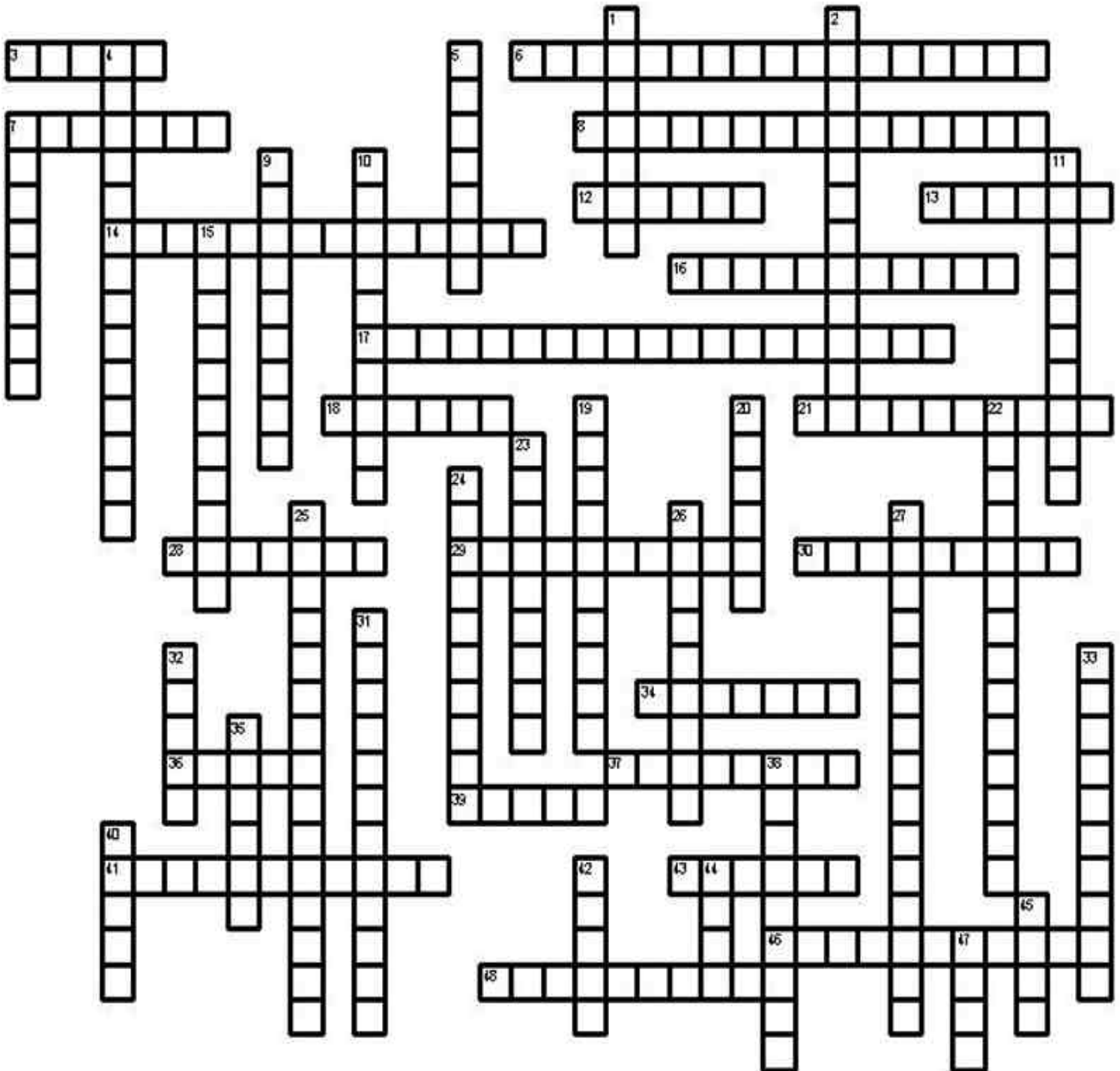
**Get the EZ solution today!**

AI Developers, Inc.  
Phone: (888) 327-9397  
info@ez-xpert.com  
www.ez-xpert.com

# Know Your AI-Q

Welcome to PC AI's feature that allows you to test your AI-Q. This is the next in a series of crossword puzzles on the various technical categories of AI. Future puzzle topics will include robotics, LISP, AI languages, expert systems, agents and many more. The answers will appear in the next issue.

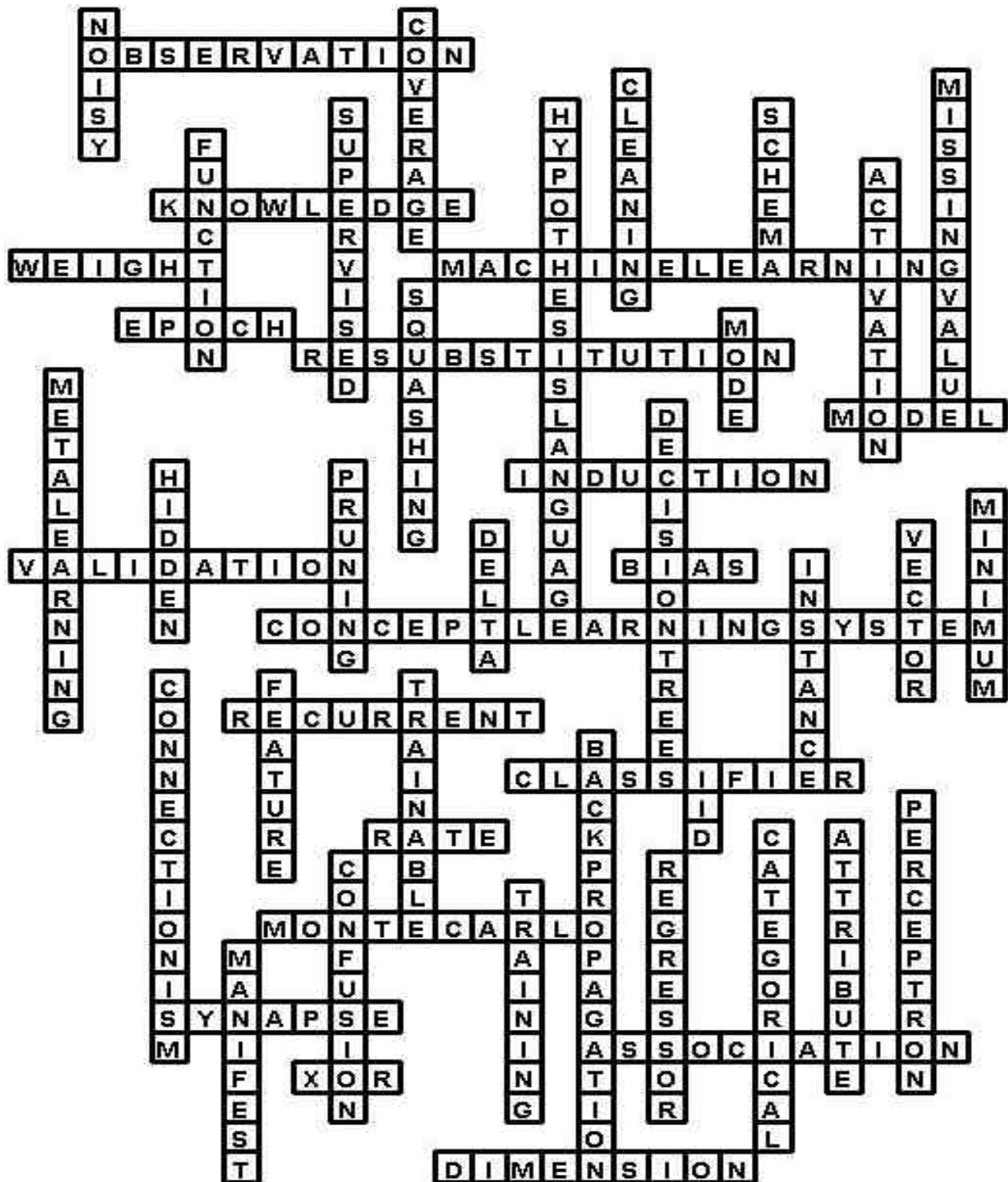
This issue's subject is *Neural Networks, Fuzzy Logic, and Speech Recognition.*



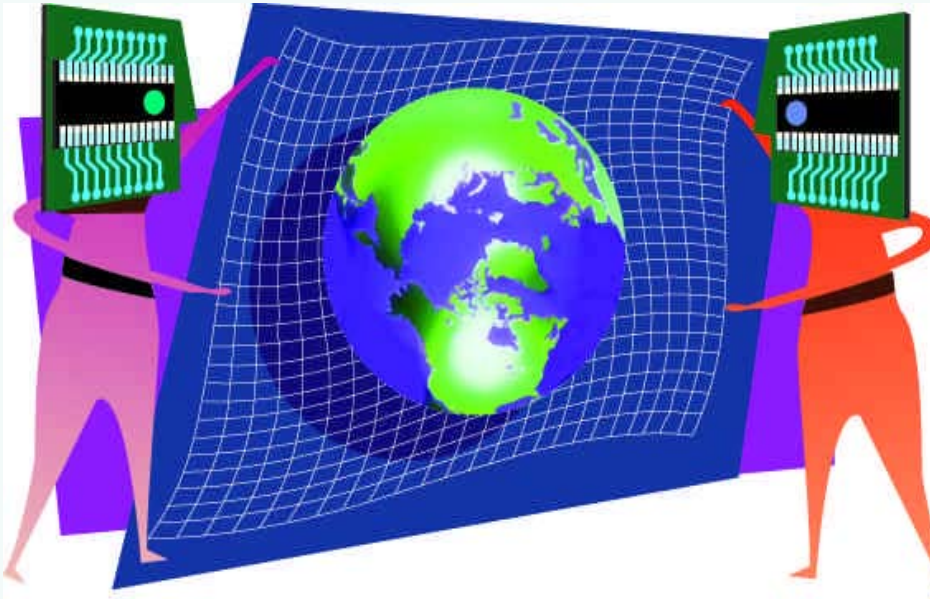


# Solution to 18.2 AI-Q Crossword Puzzle

The topic of this puzzle was *Machine Learning*



# AI and the Net



**Virtual Museum  
Education  
Assistant,  
Call Center  
Computers,  
Nose Mouse,  
Network of  
Robotic  
Telescopes,  
and more**

**Ilana Marks**

## Virtual Museum Education Assistant

The Science and Technology Museum in India has integrated a virtual education assistant called a "Cyberlady" that helps visitors learn about the exhibits and answers questions. "Cyberlady" was developed through a collaboration between the museum and the Centre for Development of Imaging Technology. When a user types in a question, the "Cyberlady" uses artificial intelligence technology to formulate a response that grabs the user's attention and provides interesting information. The "Cyberlady" responds using a synthesized voice. The "Cyberlady" learns from the conversations and stores the information gleaned from interacting with visitors so that new answers can be developed.

In addition to the "Cyberlady," the museum has also integrated virtual laboratory modules that simulate various scientific experiments. One module simulates the flame test whereby a rod is dipped in various chemicals and then is exposed to flame. The flame changes color based on what type of chemical is on the rod. Another module simulates the four-stroke engine. In this module a cross-section of an engine is observed and all of the components are identified. Other modules are in the works for the future.

<http://www.hindu.com/lj/2004/09/24/stories/2004092400500200.htm>

## AI Learns How to Cram

We all remember those late night study sessions from college where we desperately tried to fill our minds with a semester's worth of information in a few hours. Well, now there is a tool powered by AI that helps students to study smarter. The tool, called Cram101, is available at many university bookstores. Cram101 distills information from textbooks and organizes it into outline form. The outlines are printable so that students can take them to class and augment them with information from the professor's lecture.

Cram101 also creates practice tests that give students a means of diagnosing the level to which they have learned the material. Since students often have short attention spans and bore easily, the practice tests are presented in the form of a game. The tests are designed to be a self-teaching tool rather than an accurate representation of the

types of tests found in the classroom. Therefore, the answers to questions give clues to the answers of other questions. With those connections, the student has a better chance of remembering the information. Cram101 membership is

*...Thinking Software*

**LPA's** integrated suite of advanced software tools enables you to build intelligent applications both rapidly and safely.

**LPA products feature:**

- Robust and reliable run-time performance
- Support for DLLs, DDE, OLE, ODBC, TCP/IP, HTML standards
- Graphical tools and debugging aids
- Choice of delivery environment (VB, Java, Web, Delphi)

**Modules include:**

- LPA Prolog for Windows** - leading Prolog compiler system
- Flex** - popular hybrid expert system toolkit
- Agent** - distributed agent toolkit
- DataMite** - powerful data mining algorithm
- Flint** - fuzzy and probabilistic reasoning

**Logic Programming Associates Ltd**

**LPA** [www.lpa.co.uk](http://www.lpa.co.uk) [info@lpa.co.uk](mailto:info@lpa.co.uk)  
Phone: +44 (0)20 8871 2016  
Fax: +44 (0)20 8874 0449  
US Toll free: 1 800 949 7567



available for a monthly fee.  
[www.fsunews.com/vnews/display.v/ART/2004/09/23/4151f0e8c872a](http://www.fsunews.com/vnews/display.v/ART/2004/09/23/4151f0e8c872a).


## Call Center Computers

A new AI tool aims to take the frustration out of call center assistance. Once a person finally gets through to a customer service representative, often they will have to wait additional time while the operator searches through the computer to find the answer they are looking for. IBM is developing a combination speech recognition utility/search engine that listens in to the support phone calls and begins searching for the required information before the caller has finished their request. The tool works by using speech recognition to pick out key words indicating the caller's trouble. Those keywords are then entered into the call centre database, giving the operator a head start on pulling up the information.

In addition to assisting the operator in finding the correct information, the system can also alert the operator to important points that must be stressed - especially if those points constitute legal warnings. Just as the system listens to the caller's end of the conversation, it can also listen to the operator and provide on-screen reminders in order to ensure that the operator handles the call properly. While the software is still in its infancy with only a few phrases and words identified, commercial versions are in the works with a trial version scheduled to be implemented in a Dutch bank.

[www.newscientist.com/news/news.jsp?id=ns99996430](http://www.newscientist.com/news/news.jsp?id=ns99996430).

### Visit the PC AI Store



## Order Volume 16 on CD

### 6 Issues plus a bonus issue

## Volume 17 Is Now Available

## Model of a Multi-Agent System

In a joint research project from Los Alamos National Laboratory, the University of Houston and Rensselaer Polytechnic Institute, a model of a network of agents has been created. The goal is to use the model to predict the behavior of multi-agent systems - a task that can be very difficult. The model is based on the "minority game." In this game, agents make decisions and try to be in the minority when the results are revealed. The agents learn from experience what strategies are successful. In addition, the researchers created social networks amongst the agents so that information about strategy could be shared. Therefore, when making decisions a single agent will consider its own experience as well as the experiences of its neighbors. The agents strengthen connections with agents that provided useful recommendations in the past. This can be compared to human social interaction - we come to value the advice of certain people and thus turn to them

preferentially in the future for support. As a result of these connections, a leadership structure emerges. To optimize the model, the researchers are adjusting the connectivity of this leadership so that the influence of those agents is regulated. The information gleaned from this model could lead to the development of multi-agent systems suited to applications where human intervention has been required in the past. Since the agents learn from each other, the human component becomes less important.

[www.trnmag.com/Stories/2004/092204/Agent\\_model\\_yields\\_leadership\\_092204.html](http://www.trnmag.com/Stories/2004/092204/Agent_model_yields_leadership_092204.html)

## Network of Robotic Telescopes

With many astronomical events occurring in a very short span of time, it is important to be able to react quickly in order to observe them. Anyone who has watched a meteor shower knows that this is true. If you are not looking in the right place at the right time you will miss the show. However, if there are many people watching the meteor shower from many different positions then it is possible to observe more of the action.

British astronomers are hoping to take advantage of the information-capturing power of robotic telescopes by linking them together and controlling them with artificially intelligent software. The network is called RoboNet-1.0. Since the robotic telescopes in the network are located around the world, the connection provides a wider range of view as well as the ability to react quickly to interesting phenomena. Researchers hope to use this network to study gamma ray bursts. Gamma ray bursts are very intense energy bursts. The bursts are detected frequently; however the longest ones last

## RoboNet-1.0



# The Book Zone

## Applying UML

Text Mining: Predictive Methods for  
Analyzing Unstructured Information

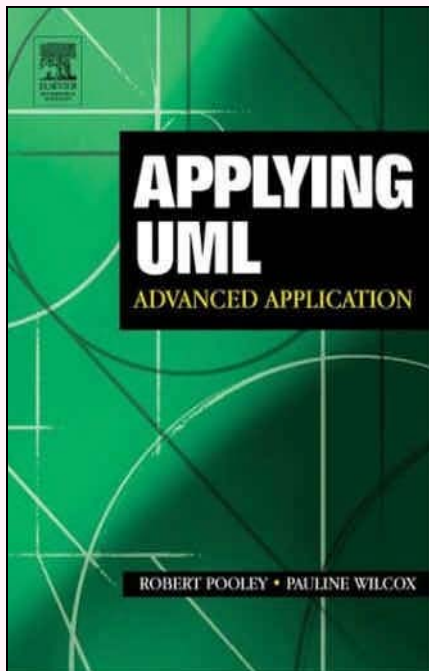
Spoken Dialogue Technology

Fuzzy Control of Queuing Systems

Data Modeling Essentials, Third Edition

Requirements Engineering

By Ilana Marks



### Applying UML by Rob Pooley and Pauline Wilcox

This book addresses issues faced by users in adopting the Unified Modeling Languages (UML) and helps them to apply it. The book covers UML in depth, including notation on profiles and extensions.

The book assumes prior experience in software engineering or business modeling, an understanding of object-oriented concepts and a basic knowledge of UML.

#### Table of Contents:

- Preface
- 1. Introduction
- 2. A Complete Example
- 3. Issues and Features
- 4. Graphics and Interaction Based Applications
- 5. Business Model
- 6. Embedded Control
- 7. Reuse
- 8. Review of Case Studies in Chapters 4,5,6, and 7
- 9. The Need for Methodologies
- 10. The Capability Model
- 11. Evaluation of Methodologies
- Appendix A - UML Notation
- Appendix B - UML Semantics
- Appendix C - Code Generation and Round Trip Engineering
- References
- Index

**Applying UML by Rob Pooley and Pauline Wilcox, November 2004, Morgan Kaufmann Publishers, ISBN: 0-7506-5683-2, Pages: 224.**

For more info:

<http://books.elsevier.com/us/mk/us/subindex.asp?isbn=0750656832&country=United+States&community=mk&ref=&mscssid=WN7TTGDMB8T8LG87CEADJGF0MGDC5R7>

### Text Mining: Predictive Methods for Analyzing Unstructured Information by Sholom M. Weiss, Nitin Indurkha, Tong Zhang and Fred Damerau

Text mining allows users to find trends and patterns in text-based information. This book analyzes new and proven techniques in text mining. It discusses topics such as automated document indexing and information retrieval and search. Also included is a look at new research in text mining including information extraction and document summarization.

#### Table of Contents:

- Preface
- 1. Overview of Text Mining
- 2. From Textual Information to Numerical Vectors
- 3. Using Text for Prediction
- 4. Information Retrieval and Text Mining



5. Finding Structure in a Document Collection
6. Looking for Information in Documents
7. Case Studies
8. Emerging Directions
- Appendix: Software Notes
- References
- Author Index
- Subject Index

**Text Mining: Predictive Methods for Analyzing Unstructured Information** by Sholom M. Weiss, Nitin Indurkha, Tong Zhang and Fred Damerau, 2004, Springer Verlag, ISBN: 0-387-95433-3, Pages: 236.

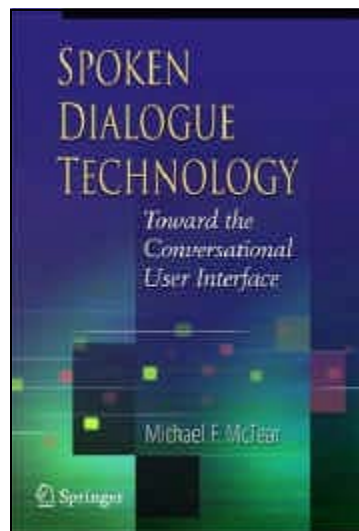
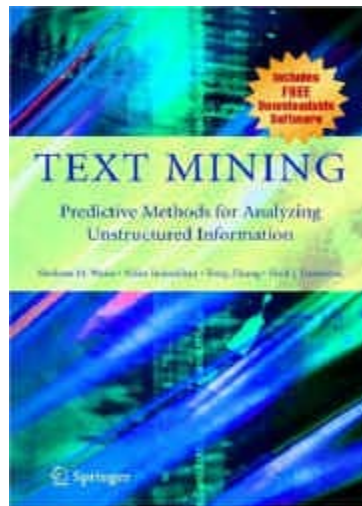
For more info:  
<http://www.springeronline.com/sgw/cda/frontpage/0,11855,4-146-22-34526885-0,00.html>

### **Spoken Dialogue Technology: Towards the Conversational User Interface** by Michael F. McTear

This book covers spoken dialogue systems, ranging from theoretical aspects to a detailed analysis of well-established methods and tools for developing spoken dialogue systems. The book enables the reader to design and test dialogue systems. Development environments and languages include the CSLU toolkit, VoiceXML, SALT, and XHTML+ voice. Research in spoken dialogue systems is presented along with theoretical issues. A dedicated web site containing supplementary materials, code and links to resources is available to readers.

#### Table of Contents:

1. What is a Spoken Dialogue System?
2. The Components of a Spoken Dialogue System
3. Describing Dialogue
4. Developing a Spoken Dialogue System: The Dialogue Engineering Lifecycle
5. Directed Dialogue Systems
6. Developing Directed Dialogue Systems using the CSLU Toolkit
7. Developing Directed Dialogue Systems using VoiceXML
8. Mixed-Initiative Dialogue Systems
9. Developing Mixed-Initiative Dialogue Systems using VoiceXML
10. Developing Mixed-Initiative Dialogue Systems using the CU Communicator System
11. Conversational Agents
12. Future Developments

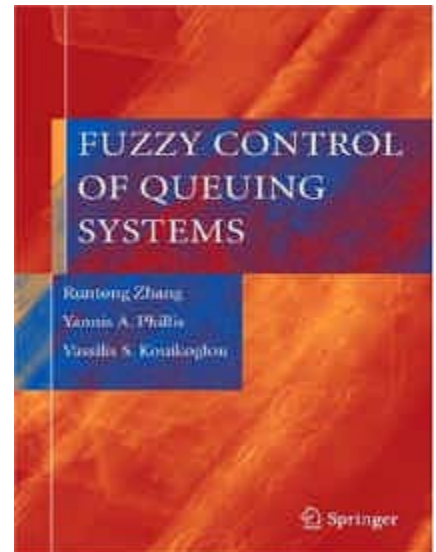


**Spoken Dialogue Technology: Towards the Conversational User Interface** by Michael F. McTear, 2004, Springer Verlag, ISBN: 1-85233-672-2, Pages: 432.

For more info:  
<http://www.springeronline.com/sgw/cda/frontpage/0,11855,4-40109-22-28205661-0,00.html>

### **Fuzzy Control of Queuing Systems** by Runtong Zhang, Yannis A. Phillis and Vassilis S. Kouikoglou

Queuing control affects manufacturing and communication networks around the world. This book discusses the use of fuzzy logic technologies to solve queuing control problems. This approach determines explicit solutions to various types of control issues in queuing systems. Included in the book are detailed case studies to demonstrate this new approach and how it differs from classical techniques. The book is directed at students, practitioners and researchers.



#### Table of Contents:

- Preface
- 1. Introduction
- 2. Fuzzy Logic
- 3. Knowledge and Fuzzy Control
- 4. Control of the Service Activities
- 5. Control of the Queue Discipline
- 6. Control of the Admission of Customers
- 7. Coordinating Multiple Control Policies
- 8. Applications of Fuzzy Queuing Control to the Internet
- 9. Closure
- Appendix: Markov Queuing Models and Simulation
- References
- Index

**Fuzzy Control of Queuing Systems** by Runtong Zhang, Yannis A. Phillis and Vassilis S. Kouikoglou, 2004, Springer Verlag, ISBN: 1-85233-824-5, Pages 175.

For more info:  
<http://www.springeronline.com/sgw/cda/frontpage/0,11855,4-40109-22-29559980-0,00.html>

**Visit the PC AI store:**

**[www.pcai.com/store](http://www.pcai.com/store)**



The vendor's address, phone number, e-mails, and URL are in a separate table after the product information table.

**Next issue:** Knowledge Representation & Management, Expert Systems, Expert System Development Tools, Machine Learning, Logic and Reasoning Systems, Intelligent Tutors

Contact PC AI for a submission form at [bg@pcai.com](mailto:bg@pcai.com)

# Buyer's Guide

Fuzzy Logic, Fuzzy SQL, Intelligent Process Control, Neural Networks, Speech Recognition, Text Mining, Web Utilities

**Special Note on System Requirements:** When *MS Windows* appears, this assumes Win 95, 98, NT, ME, 2000, XP, A Pentium Processor, 32MB RAM and 32MB Disk space available unless stated otherwise.

Fuzzy Logic			
Product Company	Description	System Requirements	Price
FLINT LPA	Fuzzy Logic toolkit with support for various treatments of uncertainty including Bayesian Updating and Certainty Factors. Includes Fuzzy Editor and support for fuzzy rule matrices.	MS Windows	Contact Vendor
Fuzzy Logic Toolbox Version 2.2 The MathWorks, Inc.	Extends MATLAB computing environment. Contains tools for the design of fuzzy logic based systems. Supports many common fuzzy logic methods including fuzzy clustering and adaptive neurofuzzy learning. Includes special GUIs, membership functions, support for AND, OR, NOT logic within rules, standard Mamdani and Sugeno-type fuzzy inference systems, automated membership function shaping. Fuzzy inference systems designed with the toolbox can be embedded in a Simulink model and C code or stand-alone executable fuzzy inference engines can be generated.	Contact Vendor	Contact Vendor
fuzzyTECH Online Edition Inform Software Corporation	Includes all features of the fuzzyTECH Professional Edition. Generates systems which can be remotely optimized on-the-fly. Fuzzy logic system is designed, tested, and generated as C code on the PC. Then, the C code is implemented and started on the target hardware. Communicates bi-directionally with the implemented code on the target system at any time via a serial cable, a network file system, or another communication link. Includes the NeuroFuzzy Module and a distribution license for the fuzzyTECH Viewer, a monitoring-only version of fuzzyTECH that can be distributed with every runtime system generated by fuzzyTECH.	Contact Vendor	Contact Vendor
fuzzyTECH Professional Edition Inform Software Corporation	Universal fuzzy logic design system for all technical application areas. Generates portable C code that can be flexibly adapted to any target hardware and deploys fuzzy logic runtime systems as DLL/ActiveX modules. fuzzyTECH's code allows for use in real-time control systems. Supports advanced fuzzy logic inference methods (compensatory operators, fuzzy associative maps, s-shape membership functions, arbitrary membership functions, etc.), and is also compatible for complex fuzzy logic applications. Allows Online debugging with the PC-based runtime modules.	Contact Vendor	Contact Vendor
Mathematica Fuzzy Logic Version 2 Wolfram Research, Inc.	Set of tools for creating, modifying and visualizing fuzzy sets and systems. Provides examples to introduce users to fuzzy logic and to show how concepts can be applied. Includes a variety of membership functions, compositions and inferencing, standard and parameterized fuzzy aggregators, fuzzy operators, fuzzy logic control, fuzzy arithmetic and more.	Mathematica 5 or later, MS Windows, Mac OS X, Linux, or Unix	Contact Vendor



Fuzzy Logic			
Product Company	Description	System Requirements	Price
Visual Rule Studio <b>Rule Machines Corp</b>	Microsoft Visual Basic ActiveX Add-on that adds a rule designer and inference engine to Visual Basic. Production Rule Language (PRL) includes fuzzy relational operators allowing developers to create rich fuzzy logic applications. Rules are authored within a rules editor which is seamlessly incorporated into the Visual Basic development environment. Rules can be organized by rule sets, and are isolated from Visual Basic code, making the rule sets easier to build, verify, and maintain. Two sample fuzzy logic projects are included with the software.	MS Windows, VB 5.0 or 6.0	Contact Vendor

Fuzzy SQL			
Product Company	Description	System Requirements	Price
dtSearch Text Retrieval Engine <b>dtSearch Corp.</b>	Lets developers add dtSearch's text retrieval to applications. Over two dozen indexed, unindexed and fielded data search options. Hit-highlighted file displays (for PDF and HTML, with images and links). Built-in HTML converters for "Office," XML, ZIP and other popular file types. Supports SQL, XML, .NET, and much more. See website for downloadable evaluations.	2 versions: dtSearch Engine for Win & .NET and dtSearch Engine for Linux.	Contact Vendor
Fuzzy Query <b>Sonalysts, Inc.</b>	A database query and data analysis tool that uses fuzzy logic to better represent the underlying semantics of data analysis questions, and ODBC to provide connectivity to a wide range of existing data stores.	Windows 95/98/NT/2000	\$59

Intelligent Process Control			
Product Company	Description	System Requirements	Price
Advisor Enterprise <b>BNH Expert Software Inc.</b>	Decision support tool manages training budgets and resources. Provides analysis for effectiveness, cost and impact of each training activity. Data resides in a central database. Managers can determine required resources to run one or multiple training programs, map where the money is spent (salaries, travel, etc.), identify what worked and why, detect problem areas, assess the impact of alternate delivery options and potential risks, evaluate build versus buy decisions, and consider multiple "what if" scenarios.	Java script browser (IE 5.5 or Netscape 6.0) for navigation bar. 32 MB of RAM, 100 MB of disk space.	\$295 for single user 1 year subscription
Amzi! Prolog + Logic Server <b>Amzi! inc</b>	Enables the integration of intelligent components with conventional applications for the addition of business rule logicbases for process control, pricing, configuration, workflow, planning, and problem solving. Access a logicbase of rules like a database. The rules are expressed in Prolog. Encapsulated as a Java Class (JSP and Servlets), C/C++ Class, .NET Class (VB, C#), Delphi Component, and DLL/SO API. Add custom Prolog predicates in Java, C/C++, C#, VB or Delphi. New Eclipse IDE with source/remote debugger.	Contact Vendor	\$0-\$1499
Common Knowledge <b>Object Connections Australia</b>	Cross-platform system for automation and maintenance of business rules throughout the application lifecycle. Comprising two parts, a rules engine and a graphical rules design studio, the system provides developers with a means to design, document, implement, and maintain their applications' business rules and logic externally in the form of Decision Tables, Trees, and scripting languages including JS, VBS, Perl and Python. Rule-sets can be changed on-the-fly without recompiling the applications, and are language and platform-independent, ideal for process control and operator guidance software development. Contact vendor for pricing on the Professional Edition.	MS Windows, Linux/Unix/Solaris, 128 MB RAM, 25-50 MB disk. Microsoft IE V4+. Supported: VB, Delphi, MS .Net, WSH, Java etc.	Contact Vendor

<b>PC AI Volume 17 now on CD - 6 issues in both <i>HTML</i> and <i>PDF</i></b> <a href="http://www.pcai.com/store">http://www.pcai.com/store</a> <b>Paid Subscribers \$14/CD, Non-Paid Subscribers \$32/CD      US Postage \$3/CD, Foreign \$5/CD</b>			
---	--	--	--

Web Utilities			
Amzi! Prolog + Logic Server <b>Amzi! inc</b>	Enables the integration of intelligent components with conventional applications for the addition of business rule logicbases for process control, pricing, configuration, workflow, planning, and problem solving. Access a logicbase of rules like a database. The rules are expressed in Prolog. Encapsulated as a Java Class (JSP and Servlets), C/C++ Class, .NET Class (VB, C#), Delphi Component, and DLL/SO API. Add custom Prolog predicates in Java, C/C++, C#, VB or Delphi. New Eclipse IDE with source/remote debugger.	Contact Vendor	\$0-\$1499
Ask.Me OnLine <b>Knowledge Broker, Inc.</b>	Available 24-hours a day the Knowledge Library contains 50,000+ solutions to FAQ's about desktop applications, operating systems and the Internet. Search engine delivers instant answers. Reduce inbound call volume and add solutions to proprietary applications. Use as a resource for a help desk. If an answer is not found, go straight to eHelpMail. For a small fee, our technicians will research your problem and e-mail the answer within 24-hours. Available to individuals, businesses and OEMs.	Access to the Internet	Contact Vendor
dtSearch Web with Spider <b>dtSearch Corp</b>	Publishes content to a Web site for searching. Over a dozen indexed and fielded data search options. Highlights hits in HTML and PDF, while displaying links and images. Converts "Office," XML, ZIP, etc. files to HTML with highlighted hits. Built-in Spider can also expand searchable database to content on a third-party site. See website for evaluations, press reviews and case studies.	Runs on IIS-based Web sites. Linux version of dtSearch Text Retrieval Engine is also available.	Contact Vendor
Visual Rule Studio <b>Rule Machines Corporation</b>	Expert system development tool based on the Production Rule Language (PRL) and Inference Engines of LEVEL5 Object adds a rule designer and inference engine to Microsoft Visual Basic. ActiveX add-on allows developers to consolidate rules in rule sets in isolation of presentation and data connectivity code. Production Rule Language (PRL) used to define the rules contends with logic processing to write, verify, and maintain rules. Customized web based expert systems can be developed from within the Microsoft Visual Studio Integrated Development Environment (IDE). Rule sets can be reused for other web based applications as well as Microsoft Windows based systems. Inference engine can use backward chaining, forward chaining, or a combination of both.	MS Windows, VB 5.0 or 6.0	Contact Vendor

Vendor Address			
Company	Address	Phone and Email	Web Address
U.S. Vendors			
Amzi! inc.	47 Redwood Road Asheville, NC 28804	828.350.0350 <i>info@amzi.com</i>	<i>www.amzi.com</i>
Alyuda Research, Inc.	1450 Frontero Ave. Los Altos, CA 94024	425.928.3570 <i>sales@alyuda.com</i>	<i>www.alyuda.com</i>
dtSearch Corp.	6852 Tulip Hill Terrace Bethesda, MD 20816	800.IT.FINDS <i>sales@dtsearch.com</i>	<i>www.dtsearch.com</i>
Gecko Systems, Inc.	1640B Highway 212, SW Conyers, GA 30094	678.413.9236 <i>info@geckosystems.com</i>	<i>www.geckosystems.com</i>
Inform Software Corporation	222 S Riverside Plz Ste 1410 Chicago, IL 60606	312.575.0578 <i>ioffice@informusa.com</i>	<i>www.fuzzytech.com</i>
ISYS Search Software	8775 E. Orchard Rd. #811 Englewood, CO 80111	800.992.4797 <i>info@isysusa.com</i>	<i>www.isysusa.com</i>
KnowledgeBroker, Inc.	PO Box 17097 Reno, NV 89521	626.441.0702 <i>sales@kbi.com</i>	<i>www.kbi.com</i>
Megaputer Intelligence, Inc.	120W, 7th Street, Suite 310 Bloomington, IN 47404	812.330.0110 <i>sales@megaputer.com</i>	<i>www.megaputer.com</i>
National Instruments	11500 N. Mopac Expwy Austin, TX 78759-3504	512.683.0100	<i>www.ni.com</i>

Vendor Address			
Company	Address	Phone and Email	Web Address
U.S. Vendors			
NeuroDimension, Inc.	1800 N. Main Street, Suite D4 Gainesville, FL 32609	352-377-5144, 800-ND-IDEAS info@nd.com	www.nd.com
NeuralWare	230 East Main Street, Suite 200 Carnegie, PA 15106-2700	412.278.6288 sales@neuralware.com	www.neuralware.com
Promised Land Technologies	195 Church Street 11th Floor New Haven, CT 06510	203.562.7335 support@promland.com	www.promland.com
Rule Machines Corporation	51745 396th Ave. Frazee, MN 56544	218.334.3960 info@rulemachines.com	www.rulemachines.com
SAS Institute Inc.	100 SAS Campus Drive Cary, NC 27513-2414	919.677.8000	www.sas.com/index.html
Sensory, Inc.	1991 Russell Avenue Santa Clara, CA 95084-2035	408.327.9000 sensorysales@sensoryinc.com	www.sensoryinc.com
Sonalysts, Inc.	215 Parkway North Waterford, CT 06385	860.326.3772 fuzzyquery@sonalysts.com	www.sonalysts.com
Sprex	1210 NE 124th St Seattle, WA 98125	206.367.7741 info@sprex.com	http://sprex.com/index.php
SPSS Inc.	233 S. Wacker Drive, 11th Floor Chicago, IL 60606-6307	312.651.3000 sales@spss.com	www.spss.com
StatSoft, Inc.	2300 E. 14th St. Tulsa, OK 74104	918.749.1119 info@statsoft.com	www.statsoft.com
Text Analysis International, Inc.	1669-2 Hollenbeck Ave. # 501 Sunnyvale, CA 94087	408.746.9932 or 877.235.6259 info@textanalysis.com	www.textanalysis.com
The MathWorks, Inc.	3 Apple Hill Drive Natick, MA 01760-2098	508.647.7000	www.mathworks.com
The Modeling Agency	PO Box 7541 The Woodlands, TX 77387	936.321.2177 info@the-modeling-agency.com	www.the-modeling-agency.com
Ward Systems Group, Inc.	5 Hillcrest Drive Frederick, MD 21703	301.662.7950 sales@wardsystems.com	www.wardsystems.com
Wolfram Research, Inc.	100 Trade Center Drive Champaign, IL 61820-7237	217.398.0700 info@wolfram.com	www.wolfram.com
Canadian Vendors			
BNH Expert Software Inc.	4000 Steinberg Street St. Laurent, QC, H4R 2G7	514.745.4010 info@bnhexpertsoft.com	www.bnhexpertsoft.com
Universal Dynamics Technologies, Inc.	100-13700 International Place Richmond, BC, V6V 2X8	604.214.3456 info@brainwave.com	www.brainwave.com
Vendors Outside North America			
KDiscovery.com	31 Simei Street 4 #02-20 Republic of Singapore 529877	(65) 3442357 info@kDiscovery.com	www.kdiscovery.com
LPA	Studio 4, RVPB, Trinity Rd London England, SW18 3SX	+44.20887.12016 info@lpa.co.uk	www.lpa.co.uk
MacSpeech, Inc.	81 Buckland Street Chippendale N.S.W. 2008	(02) 9319 2288 info@macsense.com.au	www.macspeech.com; www.macsense.com.au
Magenta Technology	33 Glasshouse Street London W1B 5DG UK	44(0)2077436450 contact@magenta-technology.com	www.magenta-technology.com
Object Connections	PO Box 562 Wallsend, Australia, 2287	61249500577 info@objectconnections.com	www.objectconnections.com
TEMIS SA	193-197 rue de Bercy 75582 Paris Cedex, France	+33(0)140044670 info@temis-group.com	www.temis-group.com
Visit PC AI at <a href="http://www.pcai.com">www.pcai.com</a>			



## ADVERTISER INDEX

Select the company name to automatically transfer to the company's home site of interest

<u>Company</u>	<u>Pages</u>
AI Developers	9, 13, 38, 42, 61
Amzi	10, 21, 28, 66
Applied Business Intelligence	75
dtSearch	8, 11, 19, 54
Exsys Inc	16, 25, 34, 55, 64
Franz Inc	4, 14, 30, 76
Intellicrafters	68
LPA	15, 26, 31, 52
Out of Toner	75
PC AI Back Issues	3
PC AI CD	6
PDC	12, 33, 57, 61
QMC	2
Statsoft	17, 56, 60
The Modeling Agency	75
WizSoft	75

**Submit your product(s)  
for the  
PC AI Buyer's Guide to  
[bg@pcai.com](mailto:bg@pcai.com)**



LEARN HOW EXPERTS  
BUILD & DEPLOY  
PREDICTIVE MODELS

Las Vegas  
Washington, DC  
San Diego  
Orlando

Guidance and Results for  
Those Who Are Data Rich,  
Yet Information Poor

VENDOR NEUTRAL &  
APPLICATION-ORIENTED  
TRAINING COURSES

**DATA MINING  
LEVELS I, II & III**

Visit Web Site for Full  
Details & Registration


(888) 742-2454  
[www.the-modeling-agency.com/pcai-special](http://www.the-modeling-agency.com/pcai-special)

**Practical data and text mining tools**




**WizWhy®**  
analyzer and predictor  
for data bases

- reveals main patterns, interesting phenomena, unexpected cases, and issues predictions



**WizDoc™**  
concept-based search engine  
for websites

- adds search capabilities to your website



**WizRule®**  
business rules detector  
for data auditing

- exposes suspected errors in the data

[www.wizsoft.com](http://www.wizsoft.com) • (914) 361-5941 • [info@wizsoft.com](mailto:info@wizsoft.com)

**WizSoft®** • Free demo is available for download

# Out Of Toner.com

hp laserjet 4000/4050 (C4127X) - \$79.00  
hp laserjet 4200 (Q1338A) - \$99.00  
hp laserjet 4300 (Q1339A) - \$105.00

**Call  
480-922-8922**

[www.OutOfToner.com](http://www.OutOfToner.com)

Canon Apple Xerox  
LEXMARK PitneyBowes  
EPSON SHARP

All trademarks and logos of the products depicted or referenced on this site are the property of their respective owners.

## Applied Business Intelligence

**Apply advanced technology to improve:**

<b>Strategic Advantage, Identify Theft, Information Technology</b>	<b>Fraud Detection, Terrorism Risk Analysis, Risk Analysis</b>
--	--

**Available technologies include:**

<b>Data Mining Data Warehousing Business Modeling Machine Learning Statistical Analysis Declarative Methodology</b>	<b>Data Architecture Business Forecasting Business Automation Decision Systems Web-based Distribution Neural Networks</b>
---	---

**Contact us directly at (602) 363-0702**

# Expand your possibilities, not your code

## with Allegro CL<sup>®</sup> and Lisp.

### **Faster Development:**

Evolve and adapt your application while it's running. Eliminate the "edit-compile-link-run-test" cycle.

### **Increased Productivity:**

Maximize development efforts by taking advantage of programmable macros, automatic memory management and crash-proof bug-trapping.

### **Webify Seamlessly:**

Deploy your application over the Internet with AllegroServe and other commercial web/app servers.

### **Easy Connectivity:**

Integrate applications with a variety of Java, XML, COM/OLE, database and CORBA tools.

### **Fully Supported:**

All licenses include one year's support. Lisp training classes and consulting also available.

**Allegro CL helps you make the most of your code. Experience for yourself why many of today's most successful companies and top researchers won't use anything else.**

**Free trial copy!**  
Download yours today at  
<http://www.franz.com/>  
[info@franz.com](mailto:info@franz.com)



AllegroServe is a trademark, and Allegro CL is a registered trademark of Franz Inc.