NetOwl™ Server

IsoQuest, Inc.
3900 Jermantown Road
Suite 400
Fairfax, VA 22030

1. Introduction

NetOwl™ Server is a powerful text analysis software product developed by IsoQuest, Inc. It supports business intelligence by providing fast and easy access to information stored on local Intranets and the global Internet. It organizes, analyzes, and summarizes data extracted by NameTag™ and any Full-Text search engine. It then presents the data for either searching or browsing. NameTag is a data extraction and indexing tool that finds proper names and other defined entities within an input text stream. NetOwl is the total application built on the NameTag core engine. NetOwl tags each desired document or Web page by person, organization, location, relationship and description giving you a browsable "back-of-the-book index." This allows for targeting exactly key information that is sought. With NetOwl it is no longer necessary to scroll through massive amounts of text to find exactly what is useful.

2. Search Functionality

NetOwl is a server-based program which generates standard CGI commands that can be executed by any standard Web browser. NetOwl's Loader ingests documents from any text or HTML-based data source, sequentially loads the documents, then passes them on to NameTag and the full-text search engine for data extraction. The full-text search engine data is stored in a proprietary database while NameTag's extracted data is stored in any ODBC-compliant database.

There are two major functional areas of the NetOwl system. The NetOwl Server system consists of:

- Loader
- Full-Text Search
- NameTag
- Database interface
- Client
-

The second NetOwl functional area comprises customization, maintenance and monitoring tools which provide support to the underlying processes of NetOwl:

- NetOwl Administrator Tool
- NetOwl Service Manager
-

Both major areas use the relational database for information storage and retrieval. This is the only link to each of the functional areas (i.e., all common information is accessed from the central database). The major functional areas never talk directly to each other or require or pass information between each other.

Loader

The Loader component of the NetOwl server collects information to process (by crawling specified sites, USENET news groups and/or locally accessible files) and then creates an index of that data in the NetOwl database.

Full-Text Search

This Full-Text Search (FTS) component accepts text documents from the loader and creates an index of all relevant words. The index thus created can later be searched by the client component to provide a list of relevant documents given a set of keywords.

Database

The Database component stores the index information generated by the Loader/NameTag combination. The database is a relational database linking instances of names to relationships, descriptions, aliases, and documents. NetOwl does not supply the database. It does, however, supply the interface to a database through an ODBC layer so that any RDBMS system can be used (e.g., Access, SQL Server, Oracle, FoxPro, Informix, Sybase).

Client

The term "client" refers to the user interface you will use to browse NetOwl's visual index, display summary information, and access information that is most salient. The client component is either a CGI-

compliant program or an HTTP server extension which accepts requests transmitted from your WWW browser and returns an HTML page with the results of the request. The client is the vehicle used to access index information in the database.

The GUI interface provides you with direct access into the NetOwl screens where you can design, submit, save, load and delete queries. It then communicates these queries to the database components and presents you with the full document text along with a glossary of the names, descriptions, relationships and words found by NameTag and the FTS engine.

NameTag

NameTag represents the key technology behind NetOwl's indexing ability. NameTag is a data extraction engine that identifies and interprets key elements of free text, particularly names of people, organizations and locations. For each document it reads, NameTag builds a signature that includes each name or concept, where it's found in the document, and how it's interpreted. For example, "Washington" can be a city, state, last name, or part of a company name, depending on the context. "Sun" can be a last name, one of several companies, or a celestial object. NameTag identifies these pieces of information, helps to interpret them and eliminate the ambiguity, and pulls in associated information such as ticker symbols for companies and variations of person names. The technology behind NameTag itself is a combination of a powerful pattern-matching technique known as finite state interpretation with a carefully-designed knowledge base.

3. Further Information

For more information contact:

IsoQuest, Inc.
3900 Jermantown Road
Suite 400
Fairfax, VA 22030
Tel.: (703) 293-2350
Fax: (703) 293-2353
e-mail: isoquest@isoquest.com
www.isoquest.com

This technical description is condensed from The Delphi Intranet Report, © 1996, The Delphi Consulting Group