

Intelligent Search Guide



 **NOVELL**®

For WINDOWS

USING ENVOY

TABLE OF CONTENTS

INDEX

SoftSolutions®

DOCUMENT MANAGEMENT SOFTWARE V 4.1

SoftSolutions is a registered trademark, and SpeedSearch and SuperSearch are trademarks of SoftSolutions Technology Corporation, a wholly-owned subsidiary of Novell, Inc.
All other brands and product names are trademarks or registered trademarks of their respective holders.

Version 4.1
© 1993-1995 Novell, Inc.
All Rights Reserved

167-000069-001—4/95

Novell, Inc.
1555 North Technology Way • Orem, Utah 84057-2399 U.S.A.
Telephone: (801)225-5000 • Fax: (801)229-1566

Table of Contents

Introduction	1
About This Manual	1
Documentation Conventions	4
Installation	5
Before Installation	5
Installation Procedures	5
Installation Results	10
Overview of SSIS	15
Accessing SSIS	15
Viewing SSIS Search Results	15
SSIS: a Natural Language Processor	16
SSIS vs. Preindexing Searched Documents	16
Features & Benefits	17
SoftSolutions and SSIS	19
Setup	21
Command Line Switches	22
EZLaunch Options	37
Implementation	37
Options Security Setup	37
Report Writer Setup	39
User Operations	41
Accessing SSIS	42
Copying	46
Searching	48
Understanding Hit Information	52
SSIS Search Criteria Window	52
Metamorph Search Results Window	52
SSIS Pattern Matching	55
Approximate Pattern Matcher	56
Word Approximation	56
Phrase Approximation	56
Numeric/Quantity Pattern Matcher	57

<i>SSIS Pattern Matching, continued</i>	
Regular EXpression (REX) Pattern Matcher	57
Word-List Pattern Matcher	58
Concept Search	58
Root Word Search	59
Set Logic	60
Search Criteria	61
Syntax	61
Qualifiers	62
Group: \(...)	63
Keyword: -	64
Mandatory Include: +	65
Minimum Set: @	67
Phrase: "..."	69
Operators	71
Boolean AND: AND (or { Space})	72
Boolean NOT: NOT (or ~)	73
Boolean OR: OR (or)	74
Numeric: #[operator]n	75
Proximity: W/n	77
REX Expression: X/ or /	78
Spelling Approximation: %n	88
Wildcard, Multiple Character: *	88
Wildcard, Single Character: ?	90
Search Examples	A-1
Messages	B-1
Quick Reference	C-1
User Operations	C-1
Pattern Matching	C-3
Search Criteria	C-6
Examples	C-14
Customer Support Services	D-1
Novell International Sales Centers	E-1
Index	

Introduction

Welcome to SoftSolutions® 4.1 Intelligent Search for Windows from Novell, Inc.

About This Manual

The *SoftSolutions Intelligent Search (SSIS) Guide* is divided into seven sections, five appendices and an index:

1. Introduction
 2. Installation
 3. Overview of SSIS
 4. Setup
 5. User Operations
 6. SSIS Pattern Matching
 7. Search Criteria
- A. Search Examples
 - B. Messages
 - C. Quick Reference
 - D. Customer Support Services
 - E. Novell International Sales Centers

The *Guide* is written for the system administrator (the person in charge of setting up and maintaining SSIS), and the *Quick Reference* appendix is written for the everyday user (a person performing daily operations in SSIS), which appendix can be photocopied freely.

NOTE

If the software was purchased as a stand-alone product, the licensee should assume both the system administrator and user roles.

Introduction

This section covers:

- documentation conventions
- product registration
- product support
- how to obtain updates and upgrades

Installation

Types of installations are:

- first-time or upgrade
- software update
- SSIS application definition creation

Overview of SSIS

The overview explains such topics as:

- SSIS processes
- features and benefits of SSIS
- SoftSolutions and SSIS interaction

Setup

The various considerations for setting up SSIS are:

- command line switches
- EZLaunch options
- SoftSolutions Report Writer

User Operations

Explained are:

- how to access SSIS
- how to copy from SSIS search results
- how to search in SSIS
- understanding hit information

SSIS Pattern Matching

This section details the four different types of pattern matching performed in SSIS:

- Approximation
- Numeric/Quantity
- Regular EXpression (REX)
- Word-List

Search Criteria

This section gives the searching syntax, explains the SSIS searching qualifiers, and explains the nine different search operators.

There are five different search *qualifiers* for SSIS. These are:

- group
- keyword
- mandatory include
- minimum set
- phrase

There are nine different search *operators* for SSIS. These are:

- Boolean AND, OR and NOT operators
- numeric (*#[operator]n*)
- proximity (*W/n*)
- REX expression (*X/ or /*)
- spelling approximation (*%n*)
- multiple (*) and single (?) wildcard character operators

Search Examples Appendix

Three examples in this appendix provide guidance on how a complex search can be set up.

Messages Appendix

This appendix provides the SSIS messages with explanations.

Quick Reference Appendix

This appendix provides quick reference for steps, searching formulas, and user operations.

Customer Support Services Appendix

This appendix explains how to obtain support for SoftSolutions SSIS.

Novell International Sales Centers Appendix

This appendix provides a complete list of Novell's International Sales Centers.

Documentation Conventions

A few documentation conventions are used to improve readability and enhance understanding.

bolded text	Field names (<i>e.g.</i> , Name field, Document Security field, Document Number field).
ALL CAPS	Window and dialog box names (<i>e.g.</i> , DOCUMENT PROFILE window, CHECK OUT dialog box); filenames and keyboard entries (<i>e.g.</i> , WP60 %P%F).
{Key Name}	Keyboard references are enclosed within <i>lenticular brackets</i> , usually as they appear on an IBM® PC keyboard. For example, the “escape” key is represented as {Esc}, and the “plus” key as {+}.
bold and <u>underline</u>	SoftSolutions <i>options</i> in the menu bar. The bolded and underlined letter represents the mnemonic used to choose an option using the keyboard. When using mnemonic letters to perform a step, the step is always preceded by pressing the [Alt] key (<i>e.g.</i> , typing “x” chooses the Search Full- <u>Text</u> option).
“monospaced text”	SoftSolutions messages (usually within quotes), and example texts.
[<i>Italicized in Brackets</i>]	“Place holders,” which are words that identify information to be supplied by the user (<i>e.g.</i> , C:[<i>Filename</i>]).

Installation

SoftSolutions Intelligent Search for Windows is accessed only by launching SSIS applications from within SoftSolutions 4.1 for Windows. These application definitions can be created automatically by the SSIS installation.

Before Installation

Before you install SSIS for Windows, be sure you have:

- *Microsoft Windows 3.1* or later already installed
- SoftSolutions 4.1 for Windows already installed
- a minimum of 3MB free hard disk space available to install SSIS program files
- determined *where* the SSIS programs will be placed (full paths, including drive letters)
- determined the *datapath* (from the contents of your SoftSolutions DATAPATH file, which may include a Dataset ID).

Installation Procedures

1. Insert the SSIS diskette labelled "Disk 1" into the file server's or workstation's floppy drive.
2. In Windows, choose **F**ile **R**un and select "INSTALL.EXE" from the floppy drive.
 - The following SSIS installation window is displayed:



3. Select the options needed based on the following information:

Install or Update Intelligent Search Programs

This option copies the SSIS program files to your network or hard drive.

Create SoftSolutions Launch Application Records

This option adds SSIS launch application definitions to your SoftSolutions dataset and creates text files, called “Response Files,” in your SSIS programs directory. Because the SSIS command line is too long to fit in application's command line in SoftSolutions, the command is placed into an ASCII text file, called a Response File. These Response Files contain the switches that allow SSIS to run.

- If you are performing a first-time installation, you should select *both* options.
 - If you have multiple datasets, and after performing an installation using *both* options, select *only* the “Create SoftSolutions Launch Application Records” option to update *each* dataset's application definitions, while sharing *one* SSIS program files location.
 - Select only the “Install or Update Intelligent Search Programs” option if your program files were inadvertently deleted or corrupted, *or* you have an update to the SSIS program files (that may be provided from time to time by SoftSolutions, but that do not require new application definitions for SSIS in SoftSolutions).
4. After selecting the desired installation options, click on the OK button.
 - The following window is displayed:

SoftSolutions Intelligent Search Installation

Please enter the following information.

Intelligent Search Programs Directory
h:\apps\ss\ssis\

SoftSolutions Programs Directory
h:\apps\ss\prog\

SoftSolutions Datapath [e.g., <datadir>:<id> or <server>:<socket>:<id>:NWB]
h:\apps\ss\data\40win

OK Cancel Help Browse

- If you selected only the “Install or Update Intelligent Search Program” option, the bottom two selections are not applicable and will appear grayed.

a. **Intelligent Search Programs Directory:**

Enter the path to the directory where the SSIS program files are to be installed.

- This may be off the SoftSolutions programs directory or any other directory. The path must include a drive letter and everyone who uses the product must use the same drive letter to reference that drive (*e.g.*, **H:\APPS\SS\SSIS**).

NOTE

Make a written note of this path to your SSIS programs directory. You will need it later when editing the BACKREF.PIF file.

b. **SoftSolutions Programs Directory:**

Enter the path to the SoftSolutions programs directory where the installation process can find the SoftSolutions SOFTSOL.CTX strings file (*e.g.*, **H:\APPS\SS\PROG**).

- This field will not be visited if you did not select the *Create SoftSolutions Launch Application Records* option from the preceding window.

c. **SoftSolutions Datapath:**

Enter the SoftSolutions datapath.

- For example, if you access the database through:

(1) **file services**, the syntax is:

[Path to Data Dir]:[Dataset ID]

e.g., **H:\APPS\SS\DATA\ : SS**

(2) **the NetWare SEM**, the syntax is:

[Server]:[Optional Socket #]:[Dataset ID]:NWB

e.g., **ADM:33912 : SS :NWB**

(3) a **UNIX SEM**, the syntax is:

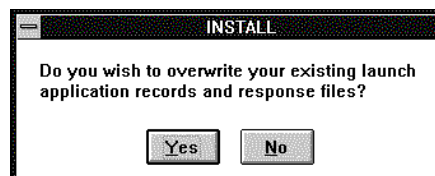
[Server Host Name]:[Socket #]:[Dataset ID]:IN

e.g., AXLE:33912:SS:IN

- The datapath entry is validated. If an improper datapath is given, an error message will be displayed and you will be returned to enter a correct path.
 - The SoftSolutions **Datapath** field will not be visited if you did not select the *Create SoftSolutions Launch Application Records* option from the preceding window.
5. If the specified SSIS programs directory does not already exist, answer the following prompt, which allows you to have the installation process create the directory for you automatically.



6. When requested, insert the remaining installation diskettes.
- If you only selected the *Create SoftSolutions Launch Application Records* option, just the first diskette will be required.
 - If you have already installed SSIS and are updating the application definitions and Response Files, *or* you already have applications defined in SoftSolutions, the following prompt will be displayed, allowing you to chose whether to overwrite the existing definitions in your SSIS subdirectory:



7. Answer Yes or No to continue.
 - Yes - Your existing application definitions and Response Files will be deleted and the defaults from this installation will be inserted.
 - No - Your existing application definitions and Response Files will be renamed and the defaults from this installation will be inserted. The application definition codes will have an X added to the end of the code, and the Response Files will have their .RSP suffixes changed to .BAK. Then, if you later desire to use your previous information, you may rename the codes and Response Files (which replaces the application definitions and Response Files created when you chose the No button).
8. *Before anyone runs **SOFTWIN.EXE***, access the Windows PIF editor.
 - a. Select the BACKREF.PIF file (from the \INT directory under the SoftSolutions programs directory).
 - b. Search for the BACKREF.EXE entry.
 - c. Edit the entry to provide the *full path* to the SSIS programs directory (where the BACKREF.EXE file was placed during installation).
 - d. Exit the PIF editor.

NOTE

As each user logs into SoftSolutions using **SOFTWIN.EXE**, the new **BACKREF.PIF** file is copied into their **WINDOWS** directory.

Installation Results

The following files are loaded into your SSIS programs directory when the *Install or Update Intelligent Search Programs* option is selected:

BACKREF.EXE	Used to index the user thesaurus.
BACKREF.PIF	PIF used to call BACKREF.EXE. This file is initially loaded into the <i>[SoftSolutions Directory]\[Programs Directory]\INT</i> directory. Before accessing SoftSolutions, the path to BACKREF.EXE needs to be manually set in BACKREF.PIF (<i>see step 8 above</i>). SoftSolutions must then be logged into with SOFTWIN.EXE, which causes the newly-edited BACKREF.PIF to be copied into the \WINDOWS directory.
BMP.DAT	For internal use of SSIS for Windows
EQUIVS	250,000-word Equivalence File.
ISWIN.EXE	SSIS executable.
ISWIN.HLP	SSIS Help file.
MORPH.DLL	Facilitates thesaurus searching.
MORPHMEM.DLL	Facilitates thesaurus searching.
OPTIONS.FIL	Example of OPTIONS file.

The following files are created in the specified directories at the indicated times:

SSIS.INT	This file is loaded into the <i>[SoftSolutions Directory]\[Programs Directory]\INT</i> directory during SSIS installation.
ISWIN.INI	This is <i>not</i> initially loaded during installation—it is created in the user's \WINDOWS directory when a Search Criteria is saved from SSIS.
SOFTSOLW.INI	This file is located in the user's \WINDOWS directory. It is modified by SOFTWIN.EXE during SSIS installation. In the [Integration Update] section, the SSIS.INT= line is added.
INTEGRAT.WKS	This file is located in the <i>[SoftSolutions Directory]\[Programs Directory]\INT</i> directory. It is updated by SOFTWIN.EXE with information from SSIS.INT. This updated entry indicates to SOFTWIN.EXE for which integrations to check its update process.

The following files are loaded into your SSIS programs directory when the *Create SoftSolutions Launch Application Records* option is selected. The *.RSP files are ASCII text files that are called Response Files in SSIS.

- CHANGES.TXT** Default CHANGES file.
- ISWINM.RSP** This Response File works with the ISWINM application definition code in SoftSolutions to launch any documents on the Mark List to SSIS. The Response File created in the SSIS programs directory is similar to the following example, but will likely have different software paths:
- ```

h:\apps\ss\ssis\iswin.exe
%r[V %n1 %v1 %n2 %v2 %n3 %v3 %n4 %v4 %n5 %v5
 %n6 %v6 %n7 %v7 %n8 %v8 %n9 %v9!!
/data=%b
/prog=%o
/user=%u
/temp=%tmp
/equivs=h:\apps\ss\ssis\equivs
/changes=h:\apps\ss\ssis\changes.txt]

```
- ISWINRW.RSP** This Response File works with the ISWINRW application definition code in SoftSolutions to launch any documents to SSIS that were selected by a designated SoftSolutions Report Writer report. The Response File created in the SSIS programs directory is similar to the following example, but will likely have different software paths:
- ```

h:\apps\ss\ssis\iswin.exe
%r[!!
/nums l=%u
/data=%b
/prog=%o
/user=%u
/temp=%tmp
/equivs=h:\apps\ss\ssis\equivs
/changes=h:\apps\ss\ssis\changes.txt]

```
- ISWINSR.RSP** This Response File works with the ISWINSR application definition code in SoftSolutions to launch documents to SSIS that were selected by the last search results. The Response File created in the SSIS programs directory is similar to the following example, but will likely have different software paths:
- ```

h:\apps\ss\ssis\iswin.exe
%r[%SR!!
/data=%b
/prog=%o
/user=%u
/temp=%tmp
/equivs=h:\apps\ss\ssis\equivs
/changes=h:\apps\ss\ssis\changes.txt
/ISearch=%sc
/parse]

```

### NOTE

The three Response File examples above show the command line switches separated by line for visual clarity. In the actual Response File, each *switch* (within the %r[] brackets) should have “white space” separating them from each other (e.g., a blank space, a CR/LF).

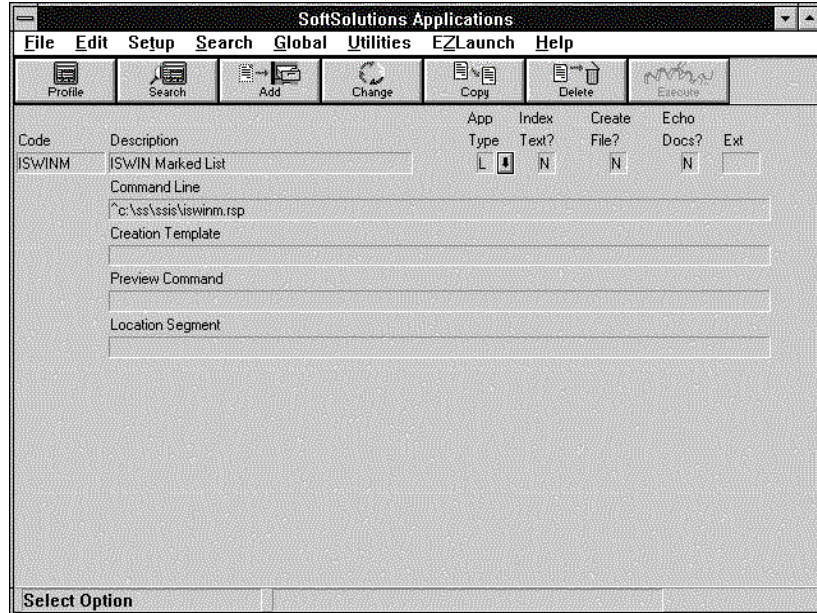
The following window illustrations show the SoftSolutions application definitions that are created during the installation process. Your directory paths may differ from the ones shown in these examples.

### NOTE

#### SoftSolutions Document & Security Server users:

If you are using the SoftSolutions Document & Security Server module, you must edit the /T=%TMP on any OAM command line for Launch or Preview so that SSIS may operate successfully. It should be changed to /T=[*Literal Path*], where [*Literal Path*] is the full path to the temporary directory (usually on the C: drive).

## Marked List Application Definition





## Report Writer Application Definition

The screenshot shows the 'SoftSolutions Applications' window with the following details:

| Code    | Description         | App Type | Index Text? | Create File? | Echo Docs? | Ext |
|---------|---------------------|----------|-------------|--------------|------------|-----|
| ISWINRW | ISWIN Report Writer | L        | N           | N            | N          |     |

Fields below the table:

- Command Line: ^c:\ss\ssis\iswinrw.rsp
- Creation Template: [Empty]
- Preview Command: [Empty]
- Location Segment: [Empty]

Buttons: Profile, Search, Add, Change, Copy, Delete, Execute.

Footer: Select Option

## Search Results Application Definition

The screenshot shows the 'SoftSolutions Applications' window with the following details:

| Code    | Description          | App Type | Index Text? | Create File? | Echo Docs? | Ext |
|---------|----------------------|----------|-------------|--------------|------------|-----|
| ISWINSR | ISWIN Search Results | L        | N           | N            | N          |     |

Fields below the table:

- Command Line: ^c:\ss\ssis\iswinsr.rsp
- Creation Template: [Empty]
- Preview Command: [Empty]
- Location Segment: [Empty]

Buttons: Profile, Search, Add, Change, Copy, Delete, Execute.

Footer: Select Option



# Overview of SSIS

SSIS is a powerful add-on module to SoftSolutions. Unlike SoftSolutions, which uses SpeedSearch™ to index and perform its searches, SSIS performs sequential character searches without using an index. SSIS is a *companion* searching program for SoftSolutions. SoftSolutions *locates documents* to search and SSIS *locates, displays, and highlights information* within those documents.

SSIS is based on *Metamorph* and is developed with the *Metamorph Application Programming Interface*, a product from *Expansion Programs International, Inc.* Many of the examples in this text are borrowed from the documentation of *Metamorph*, with the consent of *Expansion Programs International, Inc.*

## **Accessing SSIS**

---

SSIS is accessed from SoftSolutions using the SoftSolutions **Launch** or **EZLaunch** options. SSIS integrates with SoftSolutions using the SoftSolutions Mark and Launch features, using Report Writer and Launch, or launching all documents contained in a Search Results.

You can use QBE searching, Full-Text searching or the SoftSolutions Report Writer to identify documents you wish SSIS to analyze. You then pass this list to SSIS with the SoftSolutions Launch feature. SSIS analyzes the documents in this list with such powerful tools as proximity search, synonym recognition, phrase matching, word root searching, concept analysis, spelling approximation, and numeric quantity recognition.

## **Viewing SSIS Search Results**

---

For fast viewing of the search results, SSIS ensures that when your search criteria is met at its *minimal* level, the result is displayed immediately, even though there may be additional words which meet your criteria. A single mouse click brings up the next hit or the next document and its first hit. You can also go back to previous hits and open any given document while viewing a hit.

## **SSIS: a Natural Language Processor** ---

SSIS is a “Natural Language Processor” that analyzes what you ask, searches your list of documents for possible answers, and displays those portions of the documents where matches exist. It uses a dictionary of 250,000 equivalences, sophisticated methodologies to parse sentences and phrases, and algorithms for numeric and text expressions analysis.

SSIS uses a vocabulary of words and their connections, constructed in a dense matrix of associations and equivalences. These associations and equivalences could be compared to a highly sophisticated thesaurus which can be *taught* the meaning of words and phrases. This matrix acts as an intelligent language filter.

SSIS' responses can be controlled to be direct or abstract in relation to your query. An SSIS query also takes into consideration how a word is used as an expression within a sentence or paragraph. You simply enter the words to be searched and SSIS performs an analysis of your search request and begins showing text for each match in the documents.

## **SSIS vs. Preindexing Searched Documents** ---

SSIS does *not* require preindexing of documents for its full-text searching—you do not have to match the desired documents with existing indexes. The fact that SSIS does not require you to preindex your documents provides the benefit of no additional overhead and *complements* the highly-optimized SoftSolutions SpeedSearch index.

Preindexing has the following disadvantages:

- adds overhead to the full-text index
- forces maintenance and administration of the index to be more time-consuming
- increases the time and complexity of index reconstruction, integrity verification, backup, and updating.

SSIS avoids these disadvantages by not preindexing documents. Rather, SoftSolutions is first used to quickly locate a group of documents, then SSIS searches deeper within those documents for the desired information.

*Traditional indexes* can track a word and its byte displacement or location in a document. These systems isolate *each word*, however, making it impractical to perform analysis or conceptual searches on the content of the original text. By searching the text *in its native form*, SSIS is able to not only search on word occurrences, but also on concepts (intersections of words or phrases and their synonyms).

## Features & Benefits

---

The following are some of the features and benefits of SSIS.

**Search Scope.** The search scope can be based on text within sentences, paragraphs, pages, Question and Answer segments, tables, or even the entire document. SSIS searches through any size block of text in which the concept might reside.

As examples, you can find all the *sentences* dealing with love and hate, all the *paragraphs* with the topic of conflict in a geographic region, *tables* referencing both air pollution and thermal factors, and *documents* with a combination of words existing anywhere in the document.

**Phrase Searches.** Consecutive words can be found as a phrase within the defined Search Scope.

**Proximity Searches.** SSIS provides proximity searching, which locates words situated within a certain distance of each other.

**Set Logic.** SSIS incorporates Set Logic, a superset of Boolean logic, by which two or more groups or lists of words are processed and the intersections and unions of the word sets are returned as hits.

**Root Word Searches.** SSIS can ignore prefixes and suffixes or consider word conjugations and variations to arrive at root forms for the search. Therefore, SSIS finds “enjoy” and “enjoyed” in a search for “enjoyment.”

**Text & Numeric Expressions.** SSIS understands standard text expressions (*e.g.*, “frame of mind,” “United States,” “vitamin B”)—even when the quotes are omitted—and the expression is treated as a single unit. Numeric or text expressions of variable-length (*e.g.*, chemical compounds) and fixed-length (*e.g.*, phone numbers, SSNs) can be defined.

**Equivalence Searches.** SSIS includes a list of 250,000 words and their equivalences to support a built-in thesaurus and list of synonyms. As opposed to simple word substitution, SSIS understands the equivalences in terms of phrasing, word forms, acronyms, and idiomatic expressions.

**User-Defined Equivalences.** The SSIS default equivalence file is loaded during installation. You can add, reassign, and remove equivalences in order to search more precisely for those concepts you need specifically for your organization. This capability lets you teach SSIS new meanings of words and phrases.

**Approximation Searches.** SSIS can take into consideration misspellings, transpositions, and “sounds like” words.

**Numeric Quantity Searches.** SSIS can search for numbers as quantities, interpreting the various methods of writing them (*e.g.*, “15,000,” “fifteen thousand,” “15000,” or “15 thousand”). Comparisons are also supported, such as *equal to*, *greater than*, or *less than*.

**Concept Searching.** The SSIS searching methods provide a powerful tool for searching on *concepts*—intersections of words or phrases *and* their synonyms. Combinations of the methods quickly narrow the number of hits while locating the exact information within documents.

**Qualitative vs. Quantitative Proximity.** SSIS redefines simple proximity searching by providing proximity based on *qualitative* concepts, rather than only the proximity based on the *quantitative* physical distance between two words.

**Control Over Search Methodologies.** SSIS permits you to control your search methodology as you use only some methods or combine several.

**Relevance Ranking.** If the SSIS search results output is directed to an SSIS report, SSIS displays hits in the order it finds them in your documents and displays information only if it is relevant. The end of the report contains a summary sorted by relevance (documents with most hits first).

You determine the relevance ranking by adding or subtracting search items, by adjusting the optional words (with Set Logic), and by editing the equivalence table to better meet your individual organization's needs. *You* control the tightness or looseness of a correlation based on these accurate and sophisticated approaches. Relevance algorithms of other software assume the relevance depends only on redundant occurrences of keywords.

**Native Format of Documents.** As with SoftSolutions, SSIS generally operates on documents in their native formats (*e.g.*, *WordPerfect*), *not* requiring transformations into ASCII or preprocessing to be searched.

**Words Highlighted Within Context.** When SSIS locates text, it highlights the search *terms* it finds and the *scope* in which it finds those terms (the scope might be a sentence, paragraph, page, etc.).

**Copying.** While viewing search results, you can block (highlight) the desired text, and copy the block of text into the clipboard, a document, or use OLE copying.

**Security of SoftSolutions Enforced.** SSIS is closely integrated with SoftSolutions allowing you to only search documents in SSIS for which you have security access in SoftSolutions. This is accomplished because SSIS analyzes only the documents it receives as a result of a SoftSolutions search, which does not reveal secured documents.

**Lexis/Nexis-Like Syntax.** Much of the common syntax for Lexis/Nexis has been implemented into SSIS.

## **SoftSolutions and SSIS**

---

SSIS is a text searching engine. SoftSolutions is a document management system. The following chart gives more detail concerning the differences between the products and how they complement each other:

| <b>SoftSolutions 4.1</b>                                                                                                                                                                        |                                                                                                                                                                                                    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Provides Document management functions (version control, security enforcement, report writer, archival, launches Windows or Unix applications, etc). Has limited text management functionality. | Does not manage document functions. Has rich text management functionality (text navigation display, saving interesting paragraphs and sentences, Regular EXpression (REX) pattern matching, etc). |
| Searches Profile information (document name, author, type, date of creation, typist, application, security, user- defined fields, etc).                                                         | Does not search Document Profile. Searches only within document text.                                                                                                                              |
| Allows Full-Text searching of documents. Presents a list of matched documents and allows viewing and searching within any one document at a time.                                               | Displays the word and phrases in context, showing previous, current, and subsequent sentences and paragraphs while moving the search back and forth among matched documents.                       |
| Does not allow Cut & Paste while viewing documents in the SoftSolutions Preview mode.                                                                                                           | Allows copying while viewing hits within documents.                                                                                                                                                |
| Searches are based on an Index (SpeedSearch). Documents can be accessed in full text after they have been indexed.                                                                              | Searches are not based on an index. Documents can be Full-Text searched even if they have not been indexed by SpeedSearch.                                                                         |
| Searches all documents in the dataset. Can even search documents on multiple datasets (wide area networks).                                                                                     | Can only search in full text those documents preselected by SoftSolutions. Cannot search simultaneously on multiple datasets.                                                                      |

| SoftSolutions 4.1                                                                                                 |                                                                                                                                                                                                                                                               |
|-------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Use exact spelling of words (or valid wild characters) in order to use them as search criteria.                   | Exact spelling of words is not always necessary since the search engine can use a percentage of approximate spelling and equivalences (synonyms) of words and phrases. Understands numeric expressions as quantities. Understands word prefixes and suffixes. |
| Supports Boolean logic.                                                                                           | Supports Boolean and set logic.                                                                                                                                                                                                                               |
| Cannot search for phrases or words in proximity.                                                                  | Can search for phrases or words in proximity.                                                                                                                                                                                                                 |
| Searches are extremely fast. Based on SpeedSearch.                                                                | Searches are fast, but based on a linear search. Performance is not as fast as SpeedSearch.                                                                                                                                                                   |
| The search scope can be the profile or the entire document.                                                       | The search scope can be line, sentence, paragraph, page, document, or any other logical delimiter which SSIS can identify with its Regular EXpression (REX) pattern matcher.                                                                                  |
| Searches only on those documents to which you have security access.                                               | Because it is based on SoftSolutions it also searches only documents to which you have security access.                                                                                                                                                       |
| Can initiate the SoftSolutions document previewer which allows searches for up to 20 characters, exactly matched. | Always displays the document in its own viewer and can also call the SoftSolutions previewer.                                                                                                                                                                 |
| Can exist as a standalone product.                                                                                | Must have SoftSolutions 4.1 to operate.                                                                                                                                                                                                                       |
| Currently available under Windows, DOS or Unix.                                                                   | Currently available under DOS and Windows. (The server can be a Unix system, but the workstation needs to be Windows or DOS-based.)                                                                                                                           |
| SoftSolutions 4.1 can be implemented in a client/server architecture.                                             | SoftSolutions SSIS operates only on the client. Documents can be stored on the server.                                                                                                                                                                        |



# Setup

*Setup* describes issues which should be understood and addressed before using SSIS. Following this overview, each issue is described in detail.

## **Command Line Switches** (p.22)

You may wish to change the installation defaults, using command line switches to customize SSIS operations. For example, you may set up several application definitions containing the /OPTIONS= switch to allow searching documents, based on different text-length Search Scopes (*e.g.*, paragraph or page, rather than only the sentence default).

## **EZLaunch Options** (p.37)

SoftSolutions 4.1 allows you to create EZLaunch options for SSIS applications that were created during the initial installation.

You might also wish to assign security to the appropriate Screen Access Groups for the applications assigned to the EZLaunch options.

## **Report Writer Definition** (p.39)

During installation, an SSIS application definition (ISWINRW) and the associated Response File (ISWINRW.RSP) were created to allow the SoftSolutions Report Writer to pass documents to SSIS. You must also create a Report Writer Definition using the Report Writer Selection Criteria capability to identify which documents to pass to SSIS.

---

## Command Line Switches

The following optional switches can be placed within the ISWINM.RSP, ISWINRW.RSP and ISWINSR.RSP Response Files:

|            |                                                                                |
|------------|--------------------------------------------------------------------------------|
| /CHANGES=  | Customize Equivalence File                                                     |
| /COUT=     | Copy to Indicated Output File's Document Number                                |
| /DATA=     | Specify Location of SoftSolutions Data Files                                   |
| /EQUIVS=   | Specify Location of Equivalence File                                           |
| /ISEARCH=  | Specify Search Criteria                                                        |
| /NUMSFILE= | Specify Document Numbers for SSIS ( <i>vs. using a Response File</i> )         |
| /NOBANNER  | Suppresses initial bitmap                                                      |
| /OPTIONS=  | Change Search Scope                                                            |
| /PARSE     | Translate /ISEARCH Data into Valid SSIS Query                                  |
| /ROUT=     | Report to Indicated Output File's Document Number                              |
| /TEMP=%TMP | Specify Temporary Location from Workstation Record ( <i>%TMP is required</i> ) |

*Unlike SSIS 1.1 for DOS, the 4.1 for Windows version does **not** permit abbreviating the switch names.*

---

### **/CHANGES=**

This switch allows you to add, remove, and reassign equivalences to customize searches for your organization. In addition, you can reassign long expressions to shorter expressions.

*The Equivalence File (EQUIVS) that is loaded during installation contains 250,000 words.*

These changes are not made to the original Equivalence File, but are made in a supplemental file you create and associate with the /CHANGES= switch. For example you could create a file called "CHANGES" to be used with this switch (*e.g., /CHANGES=[Changes Filename]*). If [*Changes Filename*] is not in the current directory, you must use the full path (*e.g., /C=F:\SOFTSOL\SSIS\CHANGES*).

If the /CHANGES= switch is not used, SSIS uses only those equivalences in the Equivalence File (EQUIVS).

### **Syntax to Add Equivalences: *word1,word2,word3...***

An equivalence is added to a word by appending a comma and then the equivalence (no space). If the word is not in the list, then add it to the list on its own line, then append a comma and the equivalence. More than one equivalence can be added.

In the italicized instance above, *word2* and *word3* are equivalences of *word1*. *Word1* may already have a set of equivalences for it, or may be new to the equivalence list. You may add any words as equivalences after *word1*. Separate each word by only a *comma* (no spaces) and press [Enter] at the completion of the line.

For example, to add the words “building” and “design” to the word “construction,” enter the following on a new line of your “CHANGES” file:

```
construction,building,design
```

Only the words that are actually listed (*not* their equivalences) are added to the equivalence list of the first word. Thus, in the above example, only the words building and design, not their equivalences, are added to construction's equivalence list.

#### **NOTE**

No line in the CHANGES file can exceed the 128-character DOS limit (only 80 characters per line are recommended.) Where more equivalences exist for a root word than can fit onto one line, make multiple entries where the root word is repeated as the first item. Also, there must be a carriage return after each line (no word-wrap allowed).

### Syntax to Remove Equivalences: *word1~word2~word3...*

The tilde (~) removes equivalences and in this example will remove *word2* and *word3* from the list of equivalences for *word1*.

You can also use the comma (Add) in conjunction with the tilde (Remove) on the same line. For example, to remove the word “bad” and add the word “unruly” to the equivalence list of the word “mean,” enter: *mean~bad,unruly*.

#### NOTE

When you remove an equivalence word, be sure to remove the root word, if it exists, as an equivalence, as well. For example, the words *Admiration* and *Admire* both exist as equivalences for the word *Love*. If you wanted to remove the word *Admiration* from the equivalence of *Love*, you must also remove the word *Admire*. If not, *Admiration* will still be found when searching for the word *Love*.

### Syntax to Reassign Equivalences: *word1=word2*

An equivalence is reassigned by specifying the word you want to reassign, and equals sign, and then the word to which it is assigned (no spaces). In the italicized instance above, *word1* is reassigned to *word2*, but only to *word2* and **not** its equivalences. Any searches for *word1* will find all occurrences of *word2*, but not *word2*'s equivalences, and not *word1*.

#### NOTE

Because *word1* is “reassigned” to *word2*, a search on *word1* does **not** find occurrences of *word1*—only occurrences of *word2*.

For example, to reassign the word “Alice” to “friendly,” enter the following on a new line of your “CHANGES” file: *Alice=friendly*. This change means that a search criteria using the word *Alice* now finds the phrase: “She is a **friendly** person.” It will not find: “Alice is my friend,” because searches for “Alice” find only occurrences of “friendly.”

You can use the comma and equals sign to combine adding and reassigning words on the same line. For example, the entry `word1=word2,word3,word4 ...` reassigns `word1` to `word2`, and adds `word3` and `word4` to the equivalence list of `word1`.

#### NOTE

As indicated, if you reassign `word1` to `word2`, any search for `word1` does not hit on occurrences of itself. You can change this result if you use a combination of the Add and the Reassign formats in the following way: `word1=word2,word1`. This ensures that a search for `word1` finds occurrences of both `word2` and `word1` itself.

## REX Expressions in the /CHANGES= Switch

You can also use REX Expressions to reassign short phrases to substitute for long phrases. For example, the expression: `$DATE=^digit{1,2}[/\]=\digit{1,2}[/\]=\digit{2,4}` on a new line in the “CHANGES” file, will allow you to enter `$DATE` on the SSIS search criteria line to find all occurrences of a date which follows the format of one or two digits, followed by a `\` or `/`, followed by one or two digits, followed by a `\` or `/`, followed by two or four digits. For example, hits would be found in the following sentences: “The letter was dated 05/03/1971” or “11\10\81 is their anniversary.”

*For syntax for REX Expressions used in an ASCII text file to be associated with the /CHANGES= switch does not include the X. For example, the normal REX Expression syntax used in the Search Criteria is X/n, but when used in a text file, you use /n.*

## Classification Codes

A classification codes can be used by appending a semicolon (;) the classification code, and the words to which the code applies. Any specified classification is carried to subsequent words until a new classification is entered.

Valid classification codes are as follows (case is significant):

|                  |                           |
|------------------|---------------------------|
| P = Pronoun      | c = Conjunction           |
| i = Interjection | m = Modifier              |
| n = Noun         | p = Preposition           |
| v = Verb         | u = Unknown or don't care |

For example, you could enter: *wish;n,pie in the sky,dream;v,yearn,long,pine* to indicate that “pie in the sky” and dream are noun equivalences of the word wish and that yearn, long, and pine are verb equivalences of the word wish.

## Chain Multiple “CHANGES” Files

You may “chain” or connect two equivalence modification files by placing the key string *;chain;* followed by the name of the equivalence source file to chain to the first. For example, if you had two files called *lawdict* and *changes* located in C:\SOFTSOL\SSIS directory, you could include the following line in the *lawdict* equivalence file: *;chain;C:\SOFTSOL\SSIS\CHANGES*. This would make both files' modifications available to the original EQUIVS file when you use the /CHANGES=*lawdict* switch in the Response Files.

## Example of Using the /CHANGES= Switch

The following example shows how to create a file, associate it with the /CHANGES= switch, and include that switch in a Response File. The following will be performed in this example:

- add the words “joy” and “laughter” to the word “fun”
  - remove the word “obscure” from the equivalence list of “dark”
  - reassign the word “Thompson” to “supervisor” (but still allow hits on the word “Thompson”)
1. Use a text editor program to create a DOS file named CHANGES (or any proper DOS *filename*) and enter the following:

```
fun, joy, laughter
dark~obscure
Thompson=supervisor, Thompson
```

#### NOTE

1. The root word or phrase is the first item on a line.
2. Subsequent words/phrases are on the same line and are prefixed with the edit commands (, ~ =).
3. Lines are 80 or fewer characters long to fit on a standard screen display.
4. Lines, words and phrases do not have any leading or trailing spaces.
5. An (ENTER) key is pressed after each line, even the last line.
6. Hyphenated words should be entered with a space instead of a hyphen.
7. An optional classification codes (n for noun, v for verb) can be used appending a semicolon (;), the classification code, and the words to which it applies. Any specified classification is carried to subsequent words until a new classification is entered.

2. Save the file with the name of “CHANGES” (or any proper DOS name) in ASCII format.
  - If a word processor is used to edit the file, the file must not be saved in the word processor's native format. It must be saved in ASCII format so it can be read by SSIS.
3. Add the `/CHANGES=[Changes Filename]` switch to the ISWINM.RSP, ISWINRW.RSP, and ISWINSR.RSP Response Files. If the changes file is not in the current directory (usually the SoftSolutions program directory), include the full path to the file (e.g., `/CHANGES=F:\SOFTSOL\SSIS\CHANGES`).

When you create your own “user dictionary” using the `/CHANGES` switch, SSIS will compile and create a `CHANGES.NDX` file whenever you use this switch with a response file. This `CHANGES.NDX` file is used in conjunction with the `EQUIVS` file list of equivalences.

### Errors in the Text File Using the `/CHANGES=` Switch

If an error is received during the program compilation of the `.NDX` file, SSIS places information about the error in an output file called `CHANGES.ERR` (or whatever the DOS filename is you are using with a `.ERR` extension). If an error exists, you will be prompted with “Error in user dictionary creation. See `F:\SOFTSOL\SSIS\CHANGES.ERR`” when you try to execute a Response File containing this `/CHANGES=[Changes Filename]` switch.

If this occurs, you should type the CHANGES.ERR file to your screen to obtain a listing of the errors that must be corrected. An example of an error message is “Line too long or missing newline: line 2” if you did not press the {ENTER} key after line two of your CHANGES file.

*This /CHANGES= switch must be included with the SHEQ launch if you want to display equivalences with changes you have made.*

## **/COUT=** \_\_\_\_\_

This specifies the Document Number of the output file to which a copy is made.

## **/DATA=** \_\_\_\_\_

Use the /DATA= switch to specify the location and identity of the SoftSolutions data files are located. SSIS uses this switch to retrieve information about the location of documents and workstation setups. Include the full path specifying the location of the SoftSolutions data files and the Dataset ID (e.g., F:\SOFTSOL\DATA\SS).

### **NOTE**

If you do not use the /DATA= switch, SSIS first looks for a subdirectory called DATA within the SoftSolutions programs directory. If it is not found, SSIS then looks for a subdirectory called DATA within the current directory. SSIS *does not* execute if it cannot find the SoftSolutions data in either of these directories and if the /DATA= switch is not used.

## **/EQUIVS=** \_\_\_\_\_

Use the /EQUIVS= switch to name the desired Equivalence File you are using with SSIS. Follow the switch with the path and filename of the Equivalence File (e.g., /EQUIVS=F:\SOFTSOL\SSIS\EQUIVS).

## **/ISEARCH=** \_\_\_\_\_

This specifies the search criteria. If nothing is entered, SSIS uses the last search criteria available from the Search Full-Text search template. %SC can be entered to specify using the last search template line, which is faster for SSIS than entering nothing. Text representing search criteria may be entered (useful when not wishing to use the last search template). When text is entered, the /PARSE switch must also be used (*see /PARSE below*).



## ***/NOBANNER*** \_\_\_\_\_

This switch suppresses an initial bitmap.

## ***/NUMSFILE=*** \_\_\_\_\_

This specifies the path and filename of a Response File containing the SoftSolutions Document Numbers to be used in the SSIS search. %U can be used to automatically include all Document Numbers associated with the last search results.

## ***/OPTIONS=*** \_\_\_\_\_

You may want to set up several applications definitions which search documents based on different Search Scopes. The default is to search by sentence. To change this default or to create other launches using different options, use the /OPTIONS= switch.

Desired changes are made in a DOS text file. During installation, OPTIONS.FIL was loaded and can be used as this text file. This file is then associated with the /OPTIONS= switch. For example, you could create a file called "OPTIONS" to be used with this switch (*e.g.*, /OPTIONS=[Options]).

This switch may contain any of the following setting modifications:

- Search Scope
- Suffix and Prefix
- Minimum Word Length
- Number of Required Intersections

After the /OPTIONS= switch enter the path and filename of a DOS text file which contains your desired settings.

You use the /OPTIONS= switch to change the default settings of SSIS by following these steps:

1. Create a DOS text file which contains your desired changes to the default SSIS settings (*e.g.*, OPTIONS).
2. Add the "/OPTIONS=[Options]" switch to the ISWINM.RSP, ISWINRW.RSP and ISWINSR.RSP Response Files.

Following are the options that can be placed in the OPTIONS.FIL file loaded during installation.

### ***EDEXP and SDEXP***

These REX Expressions can be used to indicate the ending delimiter expression (EDEXP=) and the starting delimiter expression (SDEXP=). These settings determine where each search scope ends and begins. They are your search delimiters. Once the starting delimiter expression is found, SSIS looks for the next occurrence of the ending delimiter expression. The text between the two delimiters is the search scope.

- The **EDEXP syntax** is `EDEXP=[Delimiter Setting]` (see “[Delimiter Settings]” on page 34). The **default** is a *sentence* delimiter setting.
- The **SDEXP syntax** is `SDEXP=[Delimiter Setting]` (see “[Delimiter Settings]” on page 34). The **default** is a *sentence* delimiter setting.

### ***EQPREFIX***

You may want to use this to specify the path and filename of where the equivalence file resides, although this particular option is not as useful as the /EQUIVS= command line switch within the Response Files.

- The **syntax** is `EQPREFIX=[Drive letter]:\[Path]\[Equivalence Filename]`. The **default** is none.

You must tell SSIS where the file EQUIVS is located in order for SSIS to use equivalences in its analysis. To do so, it is recommended that you use the /EQUIVS= command in the ISWINM.RSP, ISWINRW.RSP, and ISWINSR.RSP Response Files.

If you use the EQPREFIX= setting and you do not have the equivalence file on your path and you do not use the /EQUIVS= command line option, SSIS informs you upon entering SSIS that no equivalences are used. But then SSIS sees this switch and does in fact use the equivalence file.

## ***INCED***

Use this settings option to indicate if the ending delimiter should be included in the hit. To include the delimiter in the hit set INCED to 1. You *must* set inced to 1 (or TRUE, either will work) if you are using proximity delimiter settings (see “SDEXP” and “EDEXP” above). If you set INCED to 0 (or FALSE, either will work), the ending delimiter is not included in the hit.

- The **syntax** is INCED=1 or INCED=TRUE. The **default** is INCED=1.

## ***INCSD***

Use this settings option to include the starting delimiter in the hit. To do so, set INCSD to 1 (or TRUE, either will work). You *must* set INCSD to 1 if you use the proximity delimiter settings (see “SDEXP” and “EDEXP” above). If you set INCED to 0, the starting delimiter is not included in the hit.

- The **syntax** is INCSD=1 or INCSD=TRUE or INCSD=0 or INCSD=FALSE. The **default** is INCSD=0 (or FALSE).

## **NOTE**

The INCED and INCSD settings affect the Proximity (W/n) operator functionality.

## ***INTERSECTS***

Use this settings option to indicate the number of intersections you require to be located for a hit. If you change this setting to something other than the default (which is set at -1 for *all* words), you modify the mandatory number of sets to be found in a list. Be very careful in adjusting this setting.

- The **Syntax** is INTERSECTS=[Integer]. The **default** is INTERSECTS=-1.

#### **NOTE**

The default of -1 means “require at least one of each set in the search criteria.” This is just a code, and not part of a pattern. It is just the way that *Metamorph* chooses to indicate “every set is required” of all the sets in your search criteria.

If you wish to modify this setting, just put in the number of intersections. When you use the “@n” Minimum Set Search Qualifier, SSIS decrements what it sends to the *Metamorph* search engine by 1. Therefore, if you indicate @2 Tim Jan Tom, SSIS interprets this as a request that you want two of the set of words found. But when SSIS sends this request to the *Metamorph* search engine, it takes off one of the @n numbers.

In this example, *Metamorph* receives the request as intersects=1 which it interprets as find *one* intersection of all the sets in your search criteria. In other words, find at least two sets.

#### **MINWORDLEN**

Use this settings option to set the minimum remaining length of a word as SSIS strips its prefixes and/or suffixes. This allows defining some guideline about where to stop taking away portions of words in order to avoid having no words left as it searches for word roots. It tells where to stop stripping prefixes or suffixes off so that you still have words. (See “REBUILD” below.)

- The **syntax** is MINWORDLEN=[Integer]. The **default** is MINWORDLEN=5.

#### **PREFIX**

Use this settings option to indicate the list of possible prefixes which SSIS should recognize if you have the prefix stripping process turned ON. (See “PREFIXPROC” below.)

- The **syntax** is PREFIX=[A list of pre xes separated only by commas with no spaces]. The **default** is the default list. (See the *OPTIONS.FIL* on page 30 for this listing.)

## **PREFIXPROC**

Use this settings option to turn the prefix stripping processing on or off. If you turn on prefix stripping, SSIS strips off prefixes to find root words within your search criteria. If you turn off prefix stripping, SSIS does not strip off prefixes, and the root word is not searched in the Equivalences File.

- The **syntax** is `PREFIXPROC=TRUE` (turns prefix processing ON) *or* `PREFIXPROC=FALSE` (turns prefix processing OFF). The **default** is `PREFIXPROC=TRUE`.

## **REBUILD**

Use this settings option to turn the rebuild check on or off. If you turn on rebuild, once SSIS strips all prefixes and suffixes, the word will be rebuilt to ensure the root found was actually a word.

- The **syntax** is `REBUILD=TRUE` (turns rebuild check ON) *or* `REBUILD=FALSE` (turns rebuild check OFF). The **default** is `REBUILD=TRUE`.

## **SDEXP**

(See “EDEXP” above.)

## **SUFFIX**

Use this settings option to list the possible suffixes which SSIS should recognize if you have suffix stripping turned on.

- The **syntax** is `SUFFIX=[A list of suffixes separated only by commas and no spaces]`. The **default** is the `OPTIONS.FIL` file.

## **SUFFIXPROC**

Use this settings option to turn the suffix stripping processing on or off. If you turn on suffix stripping, SSIS strips off suffixes to find root words when performing searches. If you turn off suffix stripping, SSIS does not strip off suffixes, thus root words are not searched in the Equivalences File.

- The **syntax** is `suffixproc=TRUE` (turns suffix processing on) *or* `suffixproc=FALSE` (turns suffix processing off). The **default** is `suffixproc=TRUE`.

## **[Delimiter Settings]**

Delimiter Settings allow you to define the desired search scope for SSIS. The search scope can be by line, sentence, paragraph, page, or any other logical delimiter which SSIS can identify with its Regular EXpression (REX) pattern matcher.

When SSIS does a search it first finds part of the search criteria and then finds the starting and ending delimiters. If the complete search is not found within those delimiters then it continues its search. The search is still confined to the text within the beginning (SDEXP) and ending (EDEXP) delimiters.

Following is a list of possible delimiter settings. Variations may be made to these or completely new settings may be written. The settings may contain any ASCII characters or codes and must follow the REX expression format.

### **Sentence delimiter settings:**

```
sdexp=[^\digit\upper][.?!]\space;
edexp=[^\digit\upper][.?!]\space;
incsd=FALSE;
inced=TRUE;
```

The above settings will work for either an ASCII text file or *WordPerfect 5.1* document. In the above example, the sdexp= and edexp= are the starting and ending delimiter expression indicator. The [^\digit\upper] means to find a character that is not a digit (number) and a character that is not in upper case. The [?!] means to find either a . (period), or a ? (question mark), or an ! (exclamation point). The \space means to find any “white space,” which means any tab, carriage return, new line, blank, line feed, or space. Each setting line must end with a ; (semicolon).

### **Paragraph delimiter settings:**

```
sdexp=\x/
```