



Informatica® SSA-NAME3  
10.0.0

# Edit Rule Wizard Guide

© Copyright Informatica LLC 1993, 2018

This software and documentation contain proprietary information of Informatica LLC and are provided under a license agreement containing restrictions on use and disclosure and are also protected by copyright law. Reverse engineering of the software is prohibited. No part of this document may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording or otherwise) without prior consent of Informatica LLC. This Software may be protected by U.S. and/or international Patents and other Patents Pending.

Use, duplication, or disclosure of the Software by the U.S. Government is subject to the restrictions set forth in the applicable software license agreement and as provided in DFARS 227.7202-1(a) and 227.7702-3(a) (1995), DFARS 252.227-7013(1)(ii) (OCT 1988), FAR 12.212(a) (1995), FAR 52.227-19, or FAR 52.227-14 (ALT III), as applicable.

The information in this product or documentation is subject to change without notice. If you find any problems in this product or documentation, please report them to us in writing.

Informatica, Informatica Platform, Informatica Data Services, PowerCenter, PowerCenterRT, PowerCenter Connect, PowerCenter Data Analyzer, PowerExchange, PowerMart, Metadata Manager, Informatica Data Quality, Informatica Data Explorer, Informatica B2B Data Transformation, Informatica B2B Data Exchange Informatica On Demand, Informatica Identity Resolution, Informatica Application Information Lifecycle Management, Informatica Complex Event Processing, Ultra Messaging and Informatica Master Data Management are trademarks or registered trademarks of Informatica LLC in the United States and in jurisdictions throughout the world. All other company and product names may be trade names or trademarks of their respective owners.

Portions of this software and/or documentation are subject to copyright held by third parties, including without limitation: Copyright DataDirect Technologies. All rights reserved. Copyright © Sun Microsystems. All rights reserved. Copyright © RSA Security Inc. All Rights Reserved. Copyright © Ordinal Technology Corp. All rights reserved. Copyright © Aandacht c.v. All rights reserved. Copyright Genivia, Inc. All rights reserved. Copyright Isomorphic Software. All rights reserved. Copyright © Meta Integration Technology, Inc. All rights reserved. Copyright © Intalio. All rights reserved. Copyright © Oracle. All rights reserved. Copyright © Adobe Systems Incorporated. All rights reserved. Copyright © DataArt, Inc. All rights reserved. Copyright © ComponentSource. All rights reserved. Copyright © Microsoft Corporation. All rights reserved. Copyright © Rogue Wave Software, Inc. All rights reserved. Copyright © Teradata Corporation. All rights reserved. Copyright © Yahoo! Inc. All rights reserved. Copyright © Glyph & Cog, LLC. All rights reserved. Copyright © Thinkmap, Inc. All rights reserved. Copyright © Clearpace Software Limited. All rights reserved. Copyright © Information Builders, Inc. All rights reserved. Copyright © OSS Nokalva, Inc. All rights reserved. Copyright Edifecs, Inc. All rights reserved. Copyright Cleo Communications, Inc. All rights reserved. Copyright © International Organization for Standardization 1986. All rights reserved. Copyright © ej-technologies GmbH. All rights reserved. Copyright © Jaspersoft Corporation. All rights reserved. Copyright © International Business Machines Corporation. All rights reserved. Copyright © yWorks GmbH. All rights reserved. Copyright © Lucent Technologies. All rights reserved. Copyright (c) University of Toronto. All rights reserved. Copyright © Daniel Veillard. All rights reserved. Copyright © Unicode, Inc. Copyright IBM Corp. All rights reserved. Copyright © MicroQuill Software Publishing, Inc. All rights reserved. Copyright © PassMark Software Pty Ltd. All rights reserved. Copyright © LogiXML, Inc. All rights reserved. Copyright © 2003-2010 Lorenzi Davide, All rights reserved. Copyright © Red Hat, Inc. All rights reserved. Copyright © The Board of Trustees of the Leland Stanford Junior University. All rights reserved. Copyright © EMC Corporation. All rights reserved. Copyright © Flexera Software. All rights reserved. Copyright © Jinfonet Software. All rights reserved. Copyright © Apple Inc. All rights reserved. Copyright © Telerik Inc. All rights reserved. Copyright © BEA Systems. All rights reserved. Copyright © PDFlib GmbH. All rights reserved. Copyright © Orientation in Objects GmbH. All rights reserved. Copyright © Tanuki Software, Ltd. All rights reserved. Copyright © Ricebridge. All rights reserved. Copyright © Sencha, Inc. All rights reserved. Copyright © Scalable Systems, Inc. All rights reserved. Copyright © jqWidgets. All rights reserved. Copyright © Tableau Software, Inc. All rights reserved. Copyright © MaxMind, Inc. All Rights Reserved. Copyright © TMat Software s.r.o. All rights reserved. Copyright © MapR Technologies Inc. All rights reserved. Copyright © Amazon Corporate LLC. All rights reserved. Copyright © Highsoft. All rights reserved. Copyright © Python Software Foundation. All rights reserved. Copyright © BeOpen.com. All rights reserved. Copyright © CNRI. All rights reserved.

This product includes software developed by the Apache Software Foundation (<http://www.apache.org/>), and/or other software which is licensed under various versions of the Apache License (the "License"). You may obtain a copy of these Licenses at <http://www.apache.org/licenses/>. Unless required by applicable law or agreed to in writing, software distributed under these Licenses is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the Licenses for the specific language governing permissions and limitations under the Licenses.

This product includes software which was developed by Mozilla (<http://www.mozilla.org/>), software copyright The JBoss Group, LLC, all rights reserved; software copyright © 1999-2006 by Bruno Lowagie and Paulo Soares and other software which is licensed under various versions of the GNU Lesser General Public License Agreement, which may be found at <http://www.gnu.org/licenses/lgpl.html>. The materials are provided free of charge by Informatica, "as-is", without warranty of any kind, either express or implied, including but not limited to the implied warranties of merchantability and fitness for a particular purpose.

The product includes ACE(TM) and TAO(TM) software copyrighted by Douglas C. Schmidt and his research group at Washington University, University of California, Irvine, and Vanderbilt University, Copyright (©) 1993-2006, all rights reserved.

This product includes software developed by the OpenSSL Project for use in the OpenSSL Toolkit (copyright The OpenSSL Project. All Rights Reserved) and redistribution of this software is subject to terms available at <http://www.openssl.org> and <http://www.openssl.org/source/license.html>.

This product includes Curl software which is Copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://curl.haxx.se/docs/copyright.html>. Permission to use, copy, modify, and distribute this software for any purpose with or without fee is hereby granted, provided that the above copyright notice and this permission notice appear in all copies.

The product includes software copyright 2001-2005 (©) MetaStuff, Ltd. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://www.dom4j.org/license.html>.

The product includes software copyright © 2004-2007, The Dojo Foundation. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://dojotoolkit.org/license>.

This product includes ICU software which is copyright International Business Machines Corporation and others. All rights reserved. Permissions and limitations regarding this software are subject to terms available at <http://source.icu-project.org/repos/icu/icu/trunk/license.html>.

This product includes software copyright © 1996-2006 Per Bothner. All rights reserved. Your right to use such materials is set forth in the license which may be found at <http://www.gnu.org/software/kawa/Software-License.html>.

This product includes OSSP UUID software which is Copyright © 2002 Ralf S. Engelschall, Copyright © 2002 The OSSP Project Copyright © 2002 Cable & Wireless Deutschland. Permissions and limitations regarding this software are subject to terms available at <http://www.opensource.org/licenses/mit-license.php>.

This product includes software developed by Boost (<http://www.boost.org/>) or under the Boost software license. Permissions and limitations regarding this software are subject to terms available at [http://www.boost.org/LICENSE\\_1\\_0.txt](http://www.boost.org/LICENSE_1_0.txt).

This product includes software copyright © 1997-2007 University of Cambridge. Permissions and limitations regarding this software are subject to terms available at <http://www.pcre.org/license.txt>.

This product includes software copyright © 2007 The Eclipse Foundation. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://www.eclipse.org/org/documents/epl-v10.php> and at <http://www.eclipse.org/org/documents/edl-v10.php>.

This product includes software licensed under the terms at <http://www.tcl.tk/software/tcltk/license.html>, <http://www.bosrup.com/web/overlib/?License>, <http://www.stlport.org/doc/license.html>, <http://asm.ow2.org/license.html>, <http://www.cryptix.org/LICENSE.TXT>, <http://hsqldb.org/web/hsqLicense.html>, <http://httpunit.sourceforge.net/doc/license.html>, <http://jung.sourceforge.net/license.txt>, [http://www.gzip.org/zlib/zlib\\_license.html](http://www.gzip.org/zlib/zlib_license.html), <http://www.openldap.org/software/release/license.html>, <http://www.libssh2.org>, <http://slf4j.org/license.html>, <http://www.sente.ch/software/OpenSourceLicense.html>, <http://fusesource.com/downloads/license-agreements/fuse-message-broker-v-5-3-license-agreement>, <http://antlr.org/license.html>, <http://aopalliance.sourceforge.net/>, <http://www.bouncycastle.org/licence.html>, <http://www.jgraph.com/jgraphdownload.html>, <http://www.jcraft.com/jsch/LICENSE.txt>, [http://jotm.objectweb.org/bsd\\_license.html](http://jotm.objectweb.org/bsd_license.html), <http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231>, <http://www.slf4j.org/license.html>, <http://nanoxml.sourceforge.net/orig/copyright.html>, <http://www.json.org/license.html>, <http://forge.ow2.org/projects/javaservice/>, <http://www.postgresql.org/about/licence.html>, <http://www.sqlite.org/copyright.html>, <http://www.tcl.tk/software/tcltk/license.html>, <http://www.jaxen.org/faq.html>, <http://www.jdom.org/docs/faq.html>, <http://www.slf4j.org/license.html>, <http://www.iodbc.org/dataspace/iodbc/wiki/IODBC/License>, <http://www.keplerproject.org/md5/license.html>, <http://www.toedter.com/en/jcalendar/license.html>, <http://www.edankert.com/bounce/index.html>, <http://www.net-snmp.org/about/license.html>, <http://www.openmdx.org/#FAQ>, [http://www.php.net/license/3\\_01.txt](http://www.php.net/license/3_01.txt), <http://srp.stanford.edu/license.txt>, <http://www.schneider.com/blowfish.html>, <http://www.jmock.org/license.html>, <http://xsom.java.net>, <http://benalman.com/about/license/>, <https://github.com/CreateJS/EaselJS/blob/master/src/easeljs/display/Bitmap.js>, <http://www.h2database.com/html/license.html#summary>, <http://jsoncpp.sourceforge.net/LICENSE>, <http://jdbc.postgresql.org/license.html>, <http://protobuf.googlecode.com/svn/trunk/src/google/protobuf/descriptor.proto>, <https://github.com/rantav/hector/blob/master/LICENSE>, <http://web.mit.edu/Kerberos/krb5-current/doc/mitK5license.html>, <http://jibx.sourceforge.net/jibx-license.html>, <https://github.com/lyokato/libgeohash/blob/master/LICENSE>, <https://github.com/hjiang/jsonxx/blob/master/LICENSE>, <https://code.google.com/p/lz4/>, <https://github.com/jedisct1/libsodium/blob/master/LICENSE>, <http://one-jar.sourceforge.net/index.php?page=documents&file=license>, <https://github.com/EsotericSoftware/kryo/blob/master/license.txt>, <http://www.scala-lang.org/license.html>, <https://github.com/tinkerpop/blueprints/blob/master/LICENSE.txt>, <http://gee.cs.oswego.edu/dl/classes/EDU/oswego/cs/dl/util/concurrent/intro.html>, <https://aws.amazon.com/ssl/>, <https://github.com/twbs/bootstrap/blob/master/LICENSE>, <https://sourceforge.net/p/xmlunit/code/HEAD/tree/trunk/LICENSE.txt>, <https://github.com/documentcloud/underscore-contrib/blob/master/LICENSE>, and <https://github.com/apache/hbase/blob/master/LICENSE.txt>.

This product includes software licensed under the Academic Free License (<http://www.opensource.org/licenses/afl-3.0.php>), the Common Development and Distribution License (<http://www.opensource.org/licenses/cddl1.php>), the Common Public License (<http://www.opensource.org/licenses/cpl1.0.php>), the Sun Binary Code License Agreement Supplemental License Terms, the BSD License (<http://www.opensource.org/licenses/bsd-license.php>), the new BSD License (<http://opensource.org/licenses/BSD-3-Clause>), the MIT License (<http://www.opensource.org/licenses/mit-license.php>), the Artistic License (<http://www.opensource.org/licenses/artistic-license-1.0>) and the Initial Developer's Public License Version 1.0 (<http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/>).

This product includes software copyright © 2003-2006 Joe Walnes, 2006-2007 XStream Committers. All rights reserved. Permissions and limitations regarding this software are subject to terms available at <http://xstream.codehaus.org/license.html>. This product includes software developed by the Indiana University Extreme! Lab. For further information please visit <http://www.extreme.indiana.edu/>.

This product includes software Copyright (c) 2013 Frank Balluffi and Markus Moeller. All rights reserved. Permissions and limitations regarding this software are subject to terms of the MIT license.

See patents at <https://www.informatica.com/legal/patents.html>.

DISCLAIMER: Informatica LLC provides this documentation "as is" without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of noninfringement, merchantability, or use for a particular purpose. Informatica LLC does not warrant that this software or documentation is error free. The information provided in this software or documentation may include technical inaccuracies or typographical errors. The information in this software and documentation is subject to change at any time without notice.

## NOTICES

This Informatica product (the "Software") includes certain drivers (the "DataDirect Drivers") from DataDirect Technologies, an operating company of Progress Software Corporation ("DataDirect") which are subject to the following terms and conditions:

1. THE DATADIRECT DRIVERS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT.
2. IN NO EVENT WILL DATADIRECT OR ITS THIRD PARTY SUPPLIERS BE LIABLE TO THE END-USER CUSTOMER FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, CONSEQUENTIAL OR OTHER DAMAGES ARISING OUT OF THE USE OF THE ODBC DRIVERS, WHETHER OR NOT INFORMED OF THE POSSIBILITIES OF DAMAGES IN ADVANCE. THESE LIMITATIONS APPLY TO ALL CAUSES OF ACTION, INCLUDING, WITHOUT LIMITATION, BREACH OF CONTRACT, BREACH OF WARRANTY, NEGLIGENCE, STRICT LIABILITY, MISREPRESENTATION AND OTHER TORTS.

Publication Date: 2018-06-08

# Table of Contents

<b>Preface .....</b>	<b>5</b>
Learning About Informatica SSA-NAME3. ....	5
What Do I Read If. ....	7
Informatica Resources. ....	8
Informatica My Support Portal. ....	8
Informatica Documentation. ....	8
Informatica Product Availability Matrixes. ....	8
Informatica Web Site. ....	8
Informatica How-To Library. ....	9
Informatica Knowledge Base. ....	9
Informatica Support YouTube Channel. ....	9
Informatica Marketplace. ....	9
Informatica Velocity. ....	9
Informatica Global Customer Support. ....	9
 <b>Chapter 1: Introduction.....</b>	 <b>10</b>
 <b>Chapter 2: Starting the Edit Rule Wizard.....</b>	 <b>11</b>
 <b>Chapter 3: Working in the Edit Rule Wizard.....</b>	 <b>15</b>
About the Rule Interrogation Screen. ....	15
Searching for Rules about Individual Words or Phrases. ....	16
Importing a File of Names to Check for Rules. ....	17
Adding a New Rule. ....	22
Applying the New Rules. ....	25
 <b>Index. ....</b>	 <b>26</b>

# Preface

Welcome to the Informatica SSA-NAME3 Edit Rule Wizard Guide. This guide provides information to a business user on how to safely add certain types of Edit Rules to the Standard or Custom Population without requiring specific knowledge of SSA-NAME3 or support from a programmer or data analyst.

## Learning About Informatica SSA-NAME3

This section provides details of documentation available with the SSA-NAME3 product.

### Introduction to SSA-NAME3

Provides an overview of SSA-NAME3. It is written in a way that can be read by someone who has no prior experience of the product and wants a general overview of SSA-NAME3. It explains the problems SSA-NAME3 overcomes and provides an overview of how this is done. One chapter is dedicated to providing an overview for Application Programmers.

### Getting Started

This manual is intended to be the first technical material a new developer or designer reads before installing or using the SSA-NAME3 software, regardless of the platform or environment. Its goal is to help a new user get the software installed and produce a working prototype application that calls SSA-NAME3 and executes searches against their own data.

To achieve this it provides a "script" to follow which includes pointers to pertinent sections of the other manuals.

### Application & Database Design

This manual contains tips and techniques useful for setting up and optimizing a name search and matching application, including database issues, and illustrates best-practice techniques, common pitfalls, and strategies regarding the subject of name and address matching.

### Installation Guide

This manual provides information on how to install the SSA-NAME3 product.

## SSA-NAME3 Workbench User Guide

This is a guide to using the SSA-NAME3 Workbench - a Java GUI tool that helps a programmer understand and prototype SSA-NAME3 calls. The Workbench is also used for:

- Generating Sample Program Code;
- Executing SSA-NAME3 Calls;
- Testing different SSA-NAME3 run-time options;
- Producing debugging and support information for Informatica Corporation

**Note:** The Workbench in itself is not a search and match application. It assists the developer build a search and match application.

## API Reference

The ultimate goal of an SSA-NAME3 implementation is for application programs to be able to call SSA-NAME3's API Functions to build keys and search strategies and to compute match scores and decisions.

This manual describes a typical program process flow for building an identity search application, and also lists in detail each of the API Functions. It describes the parameters required by these functions and the information returned.

## Population Override Manager User's Guide

This is a guide to using the SSA-NAME3 Population Override Manager - a Java GUI tool that allows a trained data analyst to override some of the Standard Population rules that are supplied with the product, or provided in the form of a Custom Population. The types of rules that can be overridden using this tool are:

- Edit-list rules
- Frequency tables
- Scalar Frequency Tables
- Matching Purposes

**Note:** Use of this tool without proper training from Informatica should not be attempted, as improper use can adversely affect the reliability and performance of the search application(s).

## Edit Rule Wizard User's Guide

This is a guide to using the SSA-NAME3 Edit Rule Wizard - a Java GUI tool that helps a business user safely add certain types of Edit Rules to the Standard or Custom Population without requiring specific knowledge of SSA-NAME3 or support from a programmer or data analyst. The types of rules that can be added using this tool are:

- Discard a word or phrase when searching and matching (e.g. a new "noise" word)
- Add a new replacement word or phrase when searching and matching (e.g. a new "abbreviation", "nickname" or "acronym")
- Add a new compound name marker word

## Release Notes

The Release Notes contain information about what's new in this version of SSA-NAME3. It is also used to summarize any documentation updates as they are published.

## What Do I Read If. . .

### I am. . .

. . . a business manager

The INTRODUCTION TO SSA-NAME3 will address questions such as "Why have we got SSA-NAME3?", "What does SSA-NAME3 do"?

### I am. . .

. . . a system designer or DBA

The INTRODUCTION TO SSA-NAME3 will address questions such as "What resources are needed to implement SSA-NAME3?". The APPLICATION & DATABASE DESIGN manual will lead you through many of the design considerations of name search and matching applications.

### I am. . .

. . . installing SSA-NAME3

Before attempting to install SSA-NAME3 you should read the Getting Started document. This will describe the pre-requisites and help you plan the installation and implementation of SSA-NAME3. The actual installation steps for your platform are documented in the Installation Guide.

### I am. . .

. . . an Analyst or Application Programmer

A high-level overview is provided specifically for Application Programmers in the INTRODUCTION TO SSA-NAME3 manual. Before attempting to develop programs that interface with SSA-NAME3, you should also read the GETTING STARTED and APPLICATION & DATABASE DESIGN manuals, as well as experimenting with calls in the WORKBENCH USER GUIDE.

When developing the application program(s), use the API REFERENCE manual which describes a typical application and the Function parameters.

Working example programs that illustrate the calls to SSA-NAME3 in various languages are available by using the Sample Program button on the Workbench.

I want to know. . .

. . . what SSA-NAME3 does

The INTRODUCTION TO SSA-NAME3 manual gives an overview of what SSA-NAME3 does and how it does it.

I want to know. . .

. . . how to setup the database

Refer to the APPLICATION & DATABASE DESIGN manual for tips and techniques on configuring the database to store SSA-NAME3 Keys and optimizing it for searching and matching.

I want to know. . .

. . . how to code a search application

The INTRODUCTION TO SSA-NAME3 manual contains a specific section designed to get application programmers familiar with the concepts of developing an SSA-NAME3 search and match application.

The API REFERENCE GUIDE details the Function calls required and their parameters. The SSA-NAME3 WORKBENCH USER GUIDE shows how to generate a sample program in a variety of programming languages.

## Informatica Resources

### Informatica My Support Portal

As an Informatica customer, the first step in reaching out to Informatica is through the Informatica My Support Portal at <https://mysupport.informatica.com>. The My Support Portal is the largest online data integration collaboration platform with over 100,000 Informatica customers and partners worldwide.

As a member, you can:

- Access all of your Informatica resources in one place.
- Review your support cases.
- Search the Knowledge Base, find product documentation, access how-to documents, and watch support videos.
- Find your local Informatica User Group Network and collaborate with your peers.

### Informatica Documentation

The Informatica Documentation team makes every effort to create accurate, usable documentation. If you have questions, comments, or ideas about this documentation, contact the Informatica Documentation team through email at [infa\\_documentation@informatica.com](mailto:infa_documentation@informatica.com). We will use your feedback to improve our documentation. Let us know if we can contact you regarding your comments.

The Documentation team updates documentation as needed. To get the latest documentation for your product, navigate to Product Documentation from <https://mysupport.informatica.com>.

### Informatica Product Availability Matrixes

Product Availability Matrixes (PAMs) indicate the versions of operating systems, databases, and other types of data sources and targets that a product release supports. You can access the PAMs on the Informatica My Support Portal at <https://mysupport.informatica.com>.

### Informatica Web Site

You can access the Informatica corporate web site at <https://www.informatica.com>. The site contains information about Informatica, its background, upcoming events, and sales offices. You will also find product and partner information. The services area of the site includes important information about technical support, training and education, and implementation services.



## Informatica How-To Library

As an Informatica customer, you can access the Informatica How-To Library at <https://mysupport.informatica.com>. The How-To Library is a collection of resources to help you learn more about Informatica products and features. It includes articles and interactive demonstrations that provide solutions to common problems, compare features and behaviors, and guide you through performing specific real-world tasks.

## Informatica Knowledge Base

As an Informatica customer, you can access the Informatica Knowledge Base at <https://mysupport.informatica.com>. Use the Knowledge Base to search for documented solutions to known technical issues about Informatica products. You can also find answers to frequently asked questions, technical white papers, and technical tips. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team through email at [KB\\_Feedback@informatica.com](mailto:KB_Feedback@informatica.com).

## Informatica Support YouTube Channel

You can access the Informatica Support YouTube channel at <http://www.youtube.com/user/INFASupport>. The Informatica Support YouTube channel includes videos about solutions that guide you through performing specific tasks. If you have questions, comments, or ideas about the Informatica Support YouTube channel, contact the Support YouTube team through email at [supportvideos@informatica.com](mailto:supportvideos@informatica.com) or send a tweet to @INFASupport.

## Informatica Marketplace

The Informatica Marketplace is a forum where developers and partners can share solutions that augment, extend, or enhance data integration implementations. By leveraging any of the hundreds of solutions available on the Marketplace, you can improve your productivity and speed up time to implementation on your projects. You can access Informatica Marketplace at <http://www.informaticamarketplace.com>.

## Informatica Velocity

You can access Informatica Velocity at <https://mysupport.informatica.com>. Developed from the real-world experience of hundreds of data management projects, Informatica Velocity represents the collective knowledge of our consultants who have worked with organizations from around the world to plan, develop, deploy, and maintain successful data management solutions. If you have questions, comments, or ideas about Informatica Velocity, contact Informatica Professional Services at [ips@informatica.com](mailto:ips@informatica.com).

## Informatica Global Customer Support

You can contact a Customer Support Center by telephone or through the Online Support.

Online Support requires a user name and password. You can request a user name and password at <http://mysupport.informatica.com>.

The telephone numbers for Informatica Global Customer Support are available from the Informatica web site at <http://www.informatica.com/us/services-and-training/support-services/global-support-centers/>.

# CHAPTER 1

## Introduction

When searching and matching names, addresses and other identity data, certain rules are applied to words, phrases, numbers and codes to assist in candidate selection and matching.

Some of these rules are called Edit rules.

Each population packaged with SSA-NAME3 contains a set of standard Edit rules appropriate to its language, country, or application. Customers may occasionally find reason to apply additional rules which apply only to their data.

This is a guide to using the SSA-NAME3 Edit Rule Wizard – a Java GUI tool that helps a business user safely add certain types of Edit Rules to a Population rule set without requiring detailed knowledge of SSA-NAME3 or support from a programmer or data analyst.

The types of rules that can be added using the Edit Rule Wizard are:

- Ignore a word or phrase when searching and matching (example, a new "noise" word or phrase such as "ADDRESS NOT KNOWN")
- Add a new replacement word or phrase when searching and matching (example, a new "abbreviation", "acronym" or "nickname")
- Add a new "name separator" (example, the phrase "TRDNG AS" in the string INTELLISYNC TRDNG AS SEARCH SOFTWARE AMERICA , where TRDNG AS is an abbreviation for "Trading As" and serves to separate two names.

The Edit Rule Wizard might be used for any of the following reasons:

- To add known "local" rules to the Population rule-set
- To analyze the organization's data for potential new rules to be added
- In response to a particular search or match problem

**Note:** Standard Keys should be implemented if the Edit Rule Wizard is being used. See the person responsible for SSA-NAME3 implementation at your site for more information.

### [z/OS Limitations](#)

The Edit RuleWizard can not currently be used with an SSA-NAME3 Server that is running under IBM z/OS. However, it can be used with an SSA-NAME3 Server that is running under Unix System Services on the z/OS mainframe.

## CHAPTER 2

# Starting the Edit Rule Wizard

This chapter describes how to use the Edit Rule Wizard.

### Getting Started

Ensure SSA-NAME3 has been installed either locally or on a remote server, that the SSA-NAME3 Server has been started and that the Standard Population to be edited is available.

**Note:** The Edit Rule Wizard requires the SSA-NAME3 Server to be running.

The Edit Rule Wizard itself may best be installed by an IT person on the user's computer.

To start the Edit Rule Wizard, use the following:

- On a Windows platform, use the icon in the Informatica program group in the Start Menu. The Edit Rule Wizard entry screen will be appear.
- On an Unix environment the Edit Rule Wizard is started by setting the SSA-NAME3 client environment and executing \$SSABIN/erw.



The **Introduction to SSA-NAME3** button will launch a document that provides a short introduction to SSA-NAME3. It can be read by someone who has no prior experience of the product and wants a general overview about the SSA-NAME3. It explains the problems SSA-NAME3 is designed to overcome and the approaches it uses.

The **Documentation** button will give the user access to the documentation which includes an Introduction to SSA-NAME3 and the Edit Rule Wizard user's guide.

The **Enter Edit Rule Wizard** button launches the Edit Rule Wizard working environment.

### Connecting to SSA-NAME3

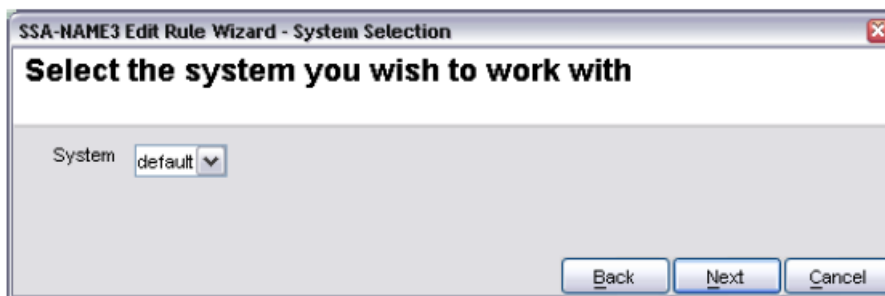
The first screen presented asks for the information required to connect to the SSA-NAME3 Server. The information is the Host Name and the Port Number and these details can normally be obtained from the person who installed SSA-NAME3.



The dialog box is titled "SSA-NAME3 Edit Rule Wizard - Connection". It has a main heading "Specify the SSA-NAME3 server to connect to". Below this, there are two input fields: "Host-Name" with the text "localhost" and "Port-Number" with the text "1665". At the bottom right, there are three buttons: "Back", "Next", and "Cancel".

### Selecting a System

The next screen asks the user to choose the System in which the Population rules reside.

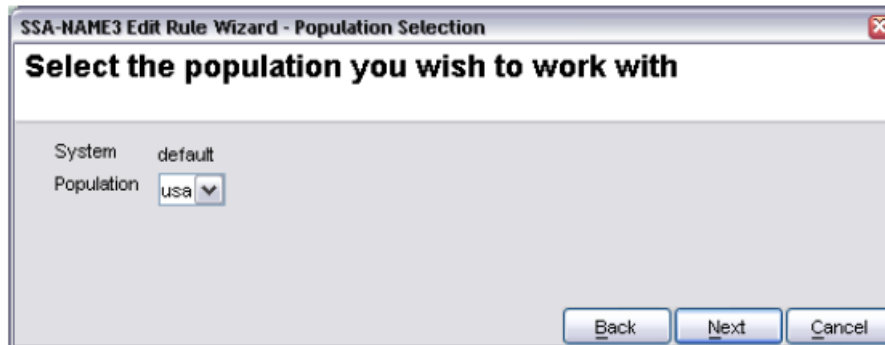


The dialog box is titled "SSA-NAME3 Edit Rule Wizard - System Selection". It has a main heading "Select the system you wish to work with". Below this, there is a "System" label followed by a drop-down menu showing "default". At the bottom right, there are three buttons: "Back", "Next", and "Cancel".

The System name is the name of the pr sub-directory where the Population files reside. The System called `default` is where the installer program copied the Standard Population files from the CD, however, other sub-directories may have been created by the person responsible for SSA-NAME3. This information should be requested from that person or team.

### Selecting a Population

The next screen asks the user to choose the Population rule-set that is to be worked with. This will typically be a country or language name, however, it could have another name if it is a Custom Population.



The dialog box is titled "SSA-NAME3 Edit Rule Wizard - Population Selection". It has a main heading "Select the population you wish to work with". Below this, there are two labels: "System" with the text "default" and "Population" with a drop-down menu showing "usa". At the bottom right, there are three buttons: "Back", "Next", and "Cancel".

### Selecting a Field

The next screen asks the user to choose the Field type to work with. The field types available in the Population selected are shown in the drop-down list.



The following field types are typically available:

#### **Person\_Name**

For adding new nicknames, salutations, honorifics, noise words or common name variations to searches that are about people.

Adding rules to the `Person_Name` field will affect searches that use `Person_Name` and matching on any Purpose that uses `Person_Name` (that is, `Person_Name`, `Individual`, `Resident`, `Household`, `Contact`, `Wide_Contact`)

#### **Organization\_Name**

For adding new abbreviations, acronyms, noise words, nicknames, name separators, or common word variations to searches that are about organizations.

Adding rules to the `Organization_Name` field will affect searches that use `Organization_Name` and matching on any Purpose that uses `Organization_Name` (that is, `Organization`, `Division`, `Corp_Entity`, `Contact`, `Wide_Contact`)

#### **Address\_Part1**

For adding new abbreviations, noise words, street types, or common word variations to searches that are about addresses. Adding rules to the `Address_Part1` field will affect searches that use `Address_Part1` and matching on any Purpose that uses `Address_Part1` (that is, `Resident`, `Contact`, `Address`, `Organization`, `Division`, `Corp_Entity`, `Household`, `Person_Name`)

#### **Id**

Allowing removal of noise from `Id` fields, or adding `Id` value replacements.

Adding rules to the `Id` field will affect matching on any Purpose that uses `Id` (that is, `Individual`, `Resident`, `Organization`, `Division`, `Person_Name`, `Corp_Entity`, `Address`, `Contact`, `Wide_Contact`)

#### **Telephone\_Number**

Allowing removal of noise from `Telephone Number` fields, or adding `Telephone Number` value replacements.

Adding rules to the `Telephone_Number` field will affect matching on any Purpose that uses `Telephone_Number` (that is, `Resident`, `Household`, `Organization`, `Division`, `Person_Name`, `Corp_Entity`, `Address`, `Contact`)

#### **Postal\_Area**

Allowing removal of noise from `Postal Area` fields, or adding `Postal Area` value replacements.

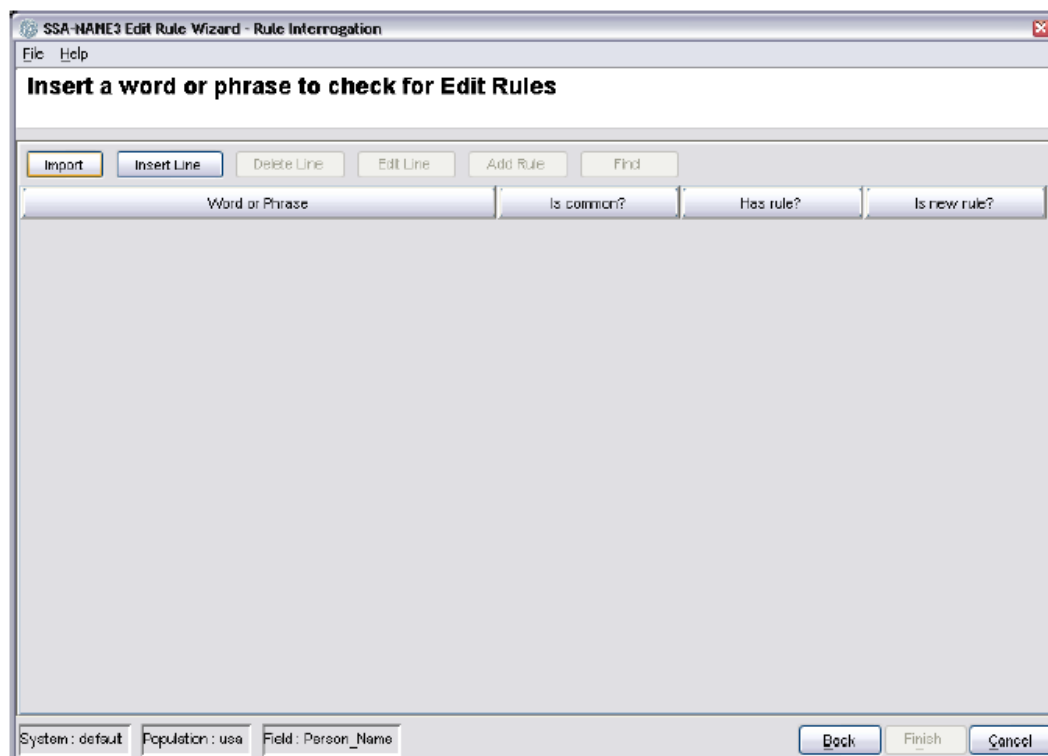
Adding rules to the `Postal_Area` field will affect matching on any Purpose that uses `Postal_Area` (that is, `Resident`, `Household`, `Organization`, `Division`, `Person_Name`, `Corp_Entity`, `Address`, `Contact`)

Click **Next** on the **Field Selection** screen will launch the **Rule Interrogation** screen.

## CHAPTER 3

# Working in the Edit Rule Wizard

After connecting to the SSA-NAME3 Server, and selecting the **System**, **Population** and **Field** that you wish to work with, the **Rule Interrogation** screen is displayed.



## About the Rule Interrogation Screen

The purpose of the Rule Interrogation screen is to allow a user to search for existing rules that have been defined about words or phrases. The search is done against the Population rule-set being used by the selected Field type and System.

If an existing rule is not found in the Population rule-set for the word or phrase entered, the user may add a new rule, or decide that no rule is necessary.

The user may either type in individual words or phrases to search for, or import a file of names to check at one time. If importing a file, an option is provided to perform a frequency analysis on the names before

checking them against the Population rule-set. This allows the user to understand the frequency of a word in their data prior to deciding if a rule is required.

A search will result in the setting of a tick or cross in the three information columns:

- Is common?
- Has rule?
- Is new rule?

### Is common?

If the **Is common?** column is ticked (✓), it indicates that when the word is used in a search, it is being optimized for performance.

**Note:** Common "noise" words in a population (example, AND, THE) appear with a cross in the **Is Common?** column and a tick in the **Has Rule?** column. This simply indicates that they are not being used in searches (that is, they are being ignored for searching and matching).

If the word is not considered common, a cross (✗) appears in this column. The **Is common?** column does not apply to phrases.

### Has rule?

If the **Has rule?** column is ticked (✓), the word or phrase already has, is part of, or is affected by, an existing rule. This may be a rule that was provided in the Standard Population delivered with the SSA-NAME3 product; a rule added in a Custom Population developed by Informatica Corporation for the organization; or a local rule added by a previous Edit Rule Wizard or Population Override Manager session.

If a cross (✗) appears in this column, the word or phrase does not have, is not part of, or is not affected by an existing rule. In this case, the **Add Rule** button will be available, and the user may choose to add a new rule for this word or phrase.

### Is new rule?

If the **Is new rule?** column is ticked (✓), a rule for the word or phrase has been added in the current Edit Rule Wizard session.

**Note:** The rule will not be available for searches until the Edit Rule Wizard session has been finished. Depending on how the person responsible for SSA-NAME3 in your organization has enabled the Edit Rule Wizard, the updated Population rule-set may then be immediately available in new searches, or may need to be moved to production and/or enabled by an IT person.

## Searching for Rules about Individual Words or Phrases

This section provides information on some Individual Words and their functions.

### Insert Line

To check for an existing rule for a single word or phrase, click the **Insert Line** button, type in the word or phrase in the text box line provided and press Enter. A search will be performed against the rules that have already been defined for this Population, including rules added in the current Edit Rule Wizard session.



The result of the search will be to fill in the columns **Is common?**, **Has rule?** and **Is new rule?**.

Word or Phrase	Is common?	Has rule?	Is new rule?
tom	✓	✓	✗

To search for another word or phrase, click the **Insert Line** button again and follow the same procedure.

#### Delete Line

To remove a line that has previously been used in a search, single-click on it to select that line (its background will change to yellow), and click on the **Delete Line** button.

To select all lines, single-click on the first line, then click **Ctrl-A**.

#### Edit Line

To reuse an existing line for another search, single-click on it to select that line (its background will change to yellow), and click on the **Edit Line** button. Alternatively, double-click on the line.

#### Find

To search for a word or phrase within the list of words or phrases in the table on the top half of the screen, click the **Find** button, and enter the word or phrase to find in the text box provided.

### RELATED TOPICS:

- [“About the Rule Interrogation Screen” on page 15](#)

## Importing a File of Names to Check for Rules

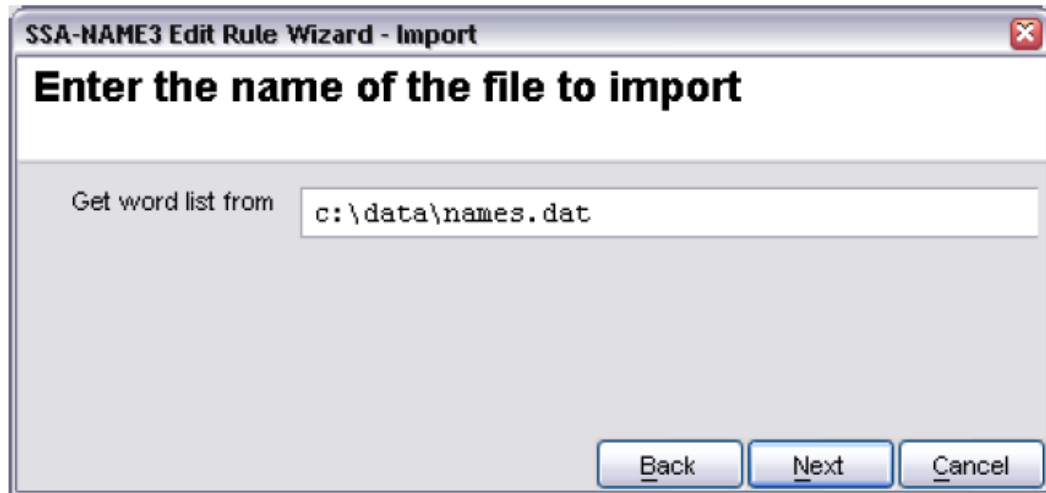
It is also possible to check an entire file of names or words and phrases against the Population rule-set.

To perform this operation, make sure the file of names or words and phrases you wish to use is on, or is copied to, the same computer where the SSA-NAME3 Server is running.

Determine and record the exact location of the file (that is, the full path name).

Because the **Edit Rule Wizard** will typically be a remote client to the SSA-NAME3 Server, a **Browse** button is not provided as part of the Import file sequence. Instead, you will be asked to enter the full name and location of the file to use.

To begin the Import process, click on the **Import** button. You will be asked to enter the full file name, including the path, of the file you wish to use.



The dialog box is titled "SSA-NAME3 Edit Rule Wizard - Import". It has a close button (X) in the top right corner. The main heading is "Enter the name of the file to import". Below this, there is a label "Get word list from" followed by a text input field containing the path "c:\data\names.dat". At the bottom right, there are three buttons: "Back", "Next", and "Cancel".

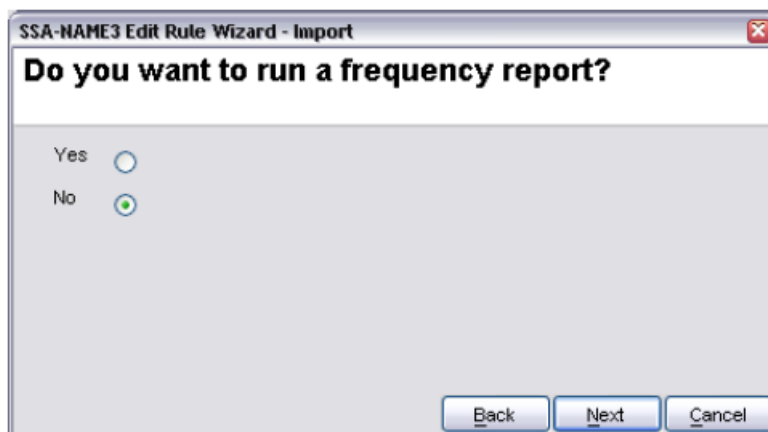
Enter the full file name, including the path, in the text box provided, then click **Next**.

#### Using a File of Words and/or Phrases

A typical reason for checking a file of words and/or phrases against the Population rule set will be if your organization has an existing list of rules about words or phrases, developed separately from Informatica Corporation. This may have been, for example, to support a legacy search application or algorithm.

The import functionality will allow you to check how many of the words and phrases in that file have rules in the SSA-NAME3 Population rule-set.

This process does not typically require a frequency analysis of the word list, therefore, when presented with the screen below, select **No**, and then click **Next** to check the file of words or phrases.



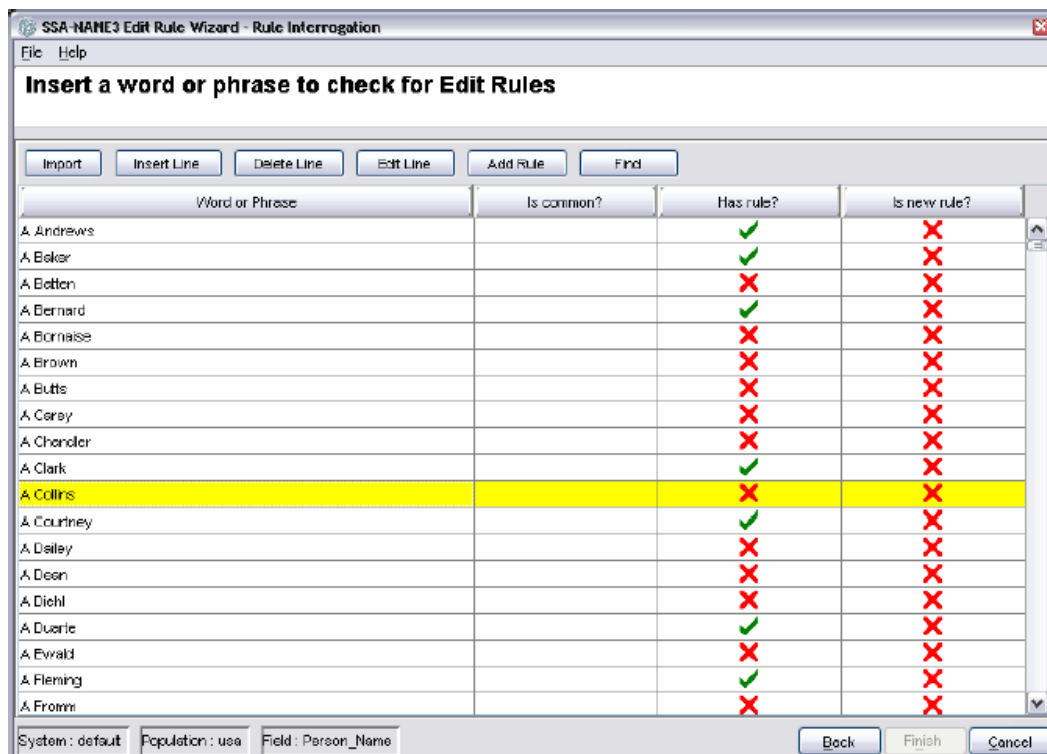
The dialog box is titled "SSA-NAME3 Edit Rule Wizard - Import". It has a close button (X) in the top right corner. The main heading is "Do you want to run a frequency report?". Below this, there are two radio button options: "Yes" and "No". The "No" option is selected, indicated by a green dot in the center of the radio button. At the bottom right, there are three buttons: "Back", "Next", and "Cancel".

After you click **Next** the following screen will be displayed. Select **Yes** to replace any currently displayed list of words or select **No** to append to the currently displayed list of words.



After you click **Next**, the **Edit Rule Wizard** will load and check all of the words in the file, and report them back to the **Rule Interrogation** screen.

For example:

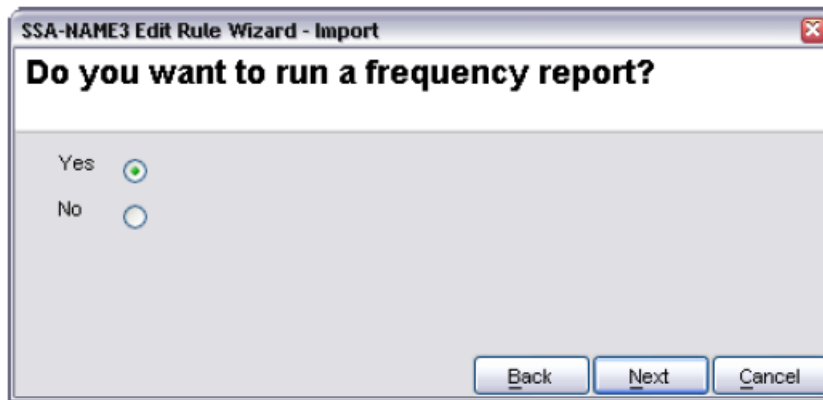


## Using a File of Names

To help a user discover the need for new rules from their own data, the Import function also allows a file of names to undergo frequency analysis, before being checked against the Population rule-set.

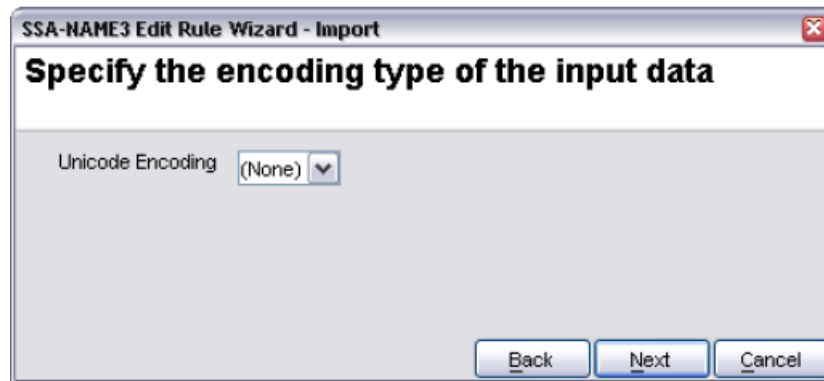
For this process to be effective, real production data should be used. Organize for a random sample of names to be extracted from the production database and written to a file. A practical number of records will be at least 50,000, or the entire file, whichever is the smaller.

Then, when presented with the screen below, select **Yes**, then click **Next** to enter the frequency report dialogue.



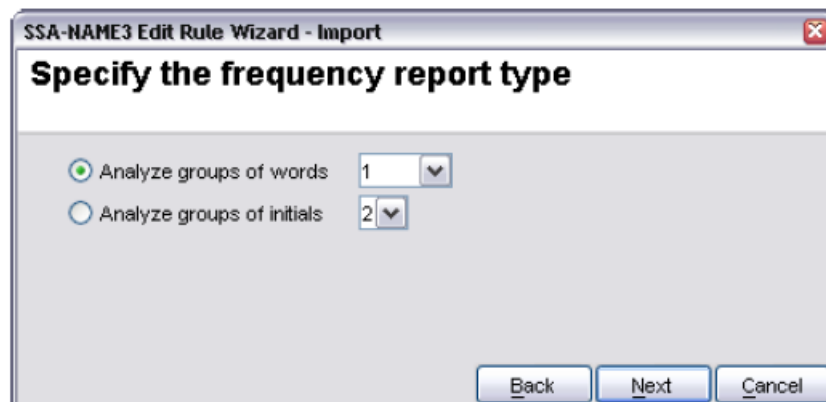
### Specifying the Unicode Encoding - if using Unicode Data

If the data is in Unicode, specify the Unicode encoding type. If the data is not Unicode, leave the Unicode encoding as "(None)".



### Specifying a Frequency Report Type

The frequency analysis of the file of names may be done on words, phrases, names or initials, by selecting options on the following screen.



To perform a frequency of analysis of single words, select: Analyze groups of words => 1

To perform a frequency analysis of phrases, specify the number of words to in the phrase, for example, for three-word phrases: Analyze groups of words => 3

To perform a frequency analysis of whole names, choose the option whole: Analyze groups of words => whole

Frequency analysis of initials can only be done on more than one initial. For example, to perform a frequency analysis of sets of three-initials, specify: Analyze groups of initials => 3

There is no option to perform a frequency analysis of initial-only names.

## Results

After you click **Next**, the **Edit Rule Wizard** will ask how many lines you wish to appear on the report. The following example shows 100 lines requested:

The image shows a dialog box titled "SSA-NAME3 Edit Rule Wizard - Import". The main heading is "The number of lines to appear on the report". Below this, there is a label "Value" followed by a text input field containing the number "100". At the bottom right, there are three buttons: "Back", "Next", and "Cancel".

Click **Next** will cause the **Edit Rule Wizard** to perform the requested frequency analysis on the file of names, check the results against the Population rule-set, and report back the 100 most frequently occurring words to the **Rule Interrogation** Screen.

For example,

The image shows a window titled "SSA-NAME3 Edit Rule Wizard - Rule Interrogation". It has a menu bar with "File" and "Help". Below the menu bar is a heading "Insert a word or phrase to check for Edit Rules". There are several buttons: "Import", "Insert Line", "Delete Line", "Edit Line", "Add Rule", and "Find". Below these buttons is a table with the following columns: "Word or Phrase", "Is common?", "Has rule?", "Is new rule?", and "Frequency". The table contains 20 rows of names. At the bottom, there are labels for "System: default", "Population: usa", and "Field: Person\_Name", along with "Back", "Finish", and "Cancel" buttons.

Word or Phrase	Is common?	Has rule?	Is new rule?	Frequency
AARON	✗	✓	✗	2
ADA	✗	✓	✗	2
ADAM	✓	✓	✗	2
AGNES	✗	✓	✗	3
AL	✓	✓	✗	21
ALAN	✓	✓	✗	12
ALBERT	✓	✓	✗	17
ALBERTA	✓	✓	✗	2
ALEX	✓	✓	✗	7
ALEXANDER	✓	✓	✗	3
ALFRED	✓	✓	✗	7
ALICE	✓	✓	✗	15
ALIDA	✓	✓	✗	2
ALLAN	✓	✓	✗	5
ALLEN	✓	✓	✗	15
ALMA	✗	✓	✗	2
ALVIN	✗	✓	✗	5
AMERICAN	✗	✗	✗	2
AMY	✓	✓	✗	3

Note the addition of the **Frequency** column on the screen. This shows the frequency of the words, phrases or names in the name file you supplied.

The default order for the report is alphabetic. To sort by the Frequency column in alternating ascending/descending order, click on the **Frequency** column heading.

SSA-NAME3 Edit Rule Wizard - Rule Interrogation

File Help

**Insert a word or phrase to check for Edit Rules**

Import Insert Line Delete Line Edit Line Add Rule Find

Word or Phrase	Is common?	Has rule?	Is new rule?	Frequency
BILL	✓	✓	✗	55
BARBARA	✓	✓	✗	43
BOB	✗	✓	✗	41
BETTY	✓	✓	✗	37
AL	✓	✓	✗	21
ANN	✓	✓	✗	20
ANTHONY	✓	✓	✗	20
ARTHUR	✓	✓	✗	19
BRENDA	✓	✓	✗	18
ALBERT	✓	✓	✗	17
BRUCE	✓	✓	✗	17
ALICE	✓	✓	✗	15
ALLEN	✓	✓	✗	15
BRIAN	✓	✓	✗	15
BEVERLY	✓	✓	✗	14
ALAN	✓	✓	✗	12
JOHNSON	✓	✗	✗	12
BEN	✓	✓	✗	11
BONNE	✓	✓	✗	11

System: default Population: usa Field: Person\_Name

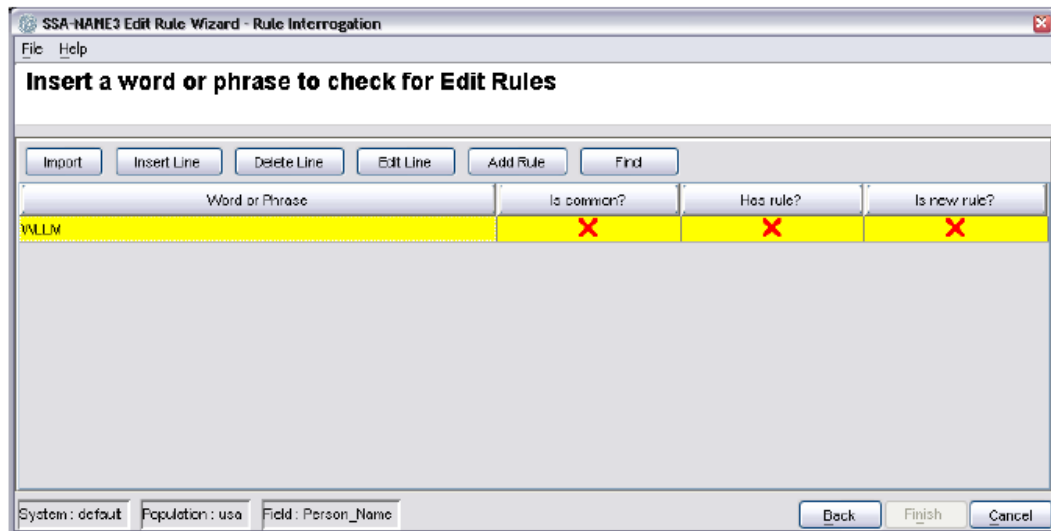
Back Finish Cancel

## Adding a New Rule

A new rule may be added for any word or phrase that does not have an existing rule defined for it. To discover whether a rule exists, use the **Rule Interrogation** screen to search for the word or phrase in question

and check the **Has Rule?** column. If it has a cross (✗), then it does not have an existing rule, and a new rule may be added.

For example, in the below screen, a search has been done in the USA Population for the abbreviation WLLM (which may be found as an abbreviation for WILLIAM). No existing rule is found, and the **Add Rule** button is available.



To add a rule for the word `WLLM`, make sure the `WLLM` line is selected (has a yellow background), then click **Add Rule**. The following screen is displayed:



For example, to cause `WLLM` to be replaced with `WILLIAM` in searches and matching, select the **Replace it with** radio button, and enter the replacement word.



When done, click the **Next** button to return to the **Rule Interrogation** screen.

Notice that a new line has been added to the new rules section at the bottom of the screen, and the `WLLM` line at the top of the screen now has a tick in the **Is new rule?** column.

SSA-NAME3 Edit Rule Wizard - Rule Interrogation

File Help

**Insert a word or phrase to check for Edit Rules**

Import Insert Line Delete Line Edit Line Add Rule Find

Word or Phrase	Is common?	Has rule?	Is new rule?
WLLM	X	X	✓

Delete Rule Edit Rule Find

Word or Phrase	Rule Type	Replace With
WLLM	Replace	WILLIAM

System : default Population : usa Field : Person\_Name Back Finish Cancel

You may continue in the current session doing more searches and adding more rules.

When complete, click the **Finish** button to cause the rules you have added during the current session to be committed to the Population rule-set.

**Note:** Depending on how the person responsible for SSA-NAME3 in your organization has enabled the **Edit Rule Wizard**, the updated Population rule-set may be immediately available in new searches, or may need to be moved to production and/or enabled by an IT person.

### Special Note Regarding Ignore Words

Be aware that when adding a new ignore word or phrase, you will no longer be able to search using that word or phrase alone unless the database keys are rebuilt.

For example, suppose you wanted to add a new ignore word `ESQUIRE` for the `Person_Name` field. Searches such as: `JOHN SMITH ESQUIRE` will still work. However, a search on: `ESQUIRE` on its own, will not return any records.

### Deleting a Rule

A rule may be deleted from the list of new rules only in the current session.

To delete a rule, select the relevant line entry on the bottom half of the screen (its background will turn yellow), then click the **Delete Rule** button.

### Editing a Rule

A rule may be edited from the list of new rules only in the current session.

To edit a rule, select the relevant line entry on the bottom half of the screen (its background will turn yellow), then click the **Edit Rule** button and make the needed changes.

### Committing Rules

Once you click the **Finish** button from the **Edit Rule Wizard**, new rules will be merged into the Population rule set and are then not available for editing through the **Edit Rule Wizard**.

If it is necessary to fix or remove a rule that has been added through the **Edit Rule Wizard**, this may be done by the analyst trained in the use of the SSA-NAME3 Population Override Manager.



## Applying the New Rules

When a user is satisfied with the new rules added, they can be committed. To do this, click the **Finish** button on the **Rule Interrogation** Screen.

The action of committing rules to a Standard or Custom Population will create what is called a Local Population. Local Populations have a file extension of `.y1p`.

When an Application program calls SSA-NAME3, SSA-NAME3 uses an order of precedence to load the Populations. If a YLP (Local Population) is present, it will be loaded. If no YLP is present, but a YCP (Custom Population) is present, it will be loaded. Otherwise the YSP (Standard Population) with that name will be loaded.

Therefore, depending on how the person responsible for SSA-NAME3 in your organization has enabled the Edit Rule Wizard, the updated Population rule-set may be immediately available in new searches, or may need to be moved to production and/or enabled by an IT person.

Even if the **Edit Rule Wizard** is enabled in production, it is likely that the new rules will not be available until the SSA-NAME3 routine is unloaded and reloaded.

# INDEX

## A

Add Rule [22](#)  
Adding rules [11](#)

## C

Committing Rules [22](#)  
Custom Population  
YCP [25](#)

## D

Delete Line [16](#)  
Delete Rule [22](#)

## E

Edit Line [16](#)  
Edit Rule [22](#)

## F

Frequency  
Rule Interrogation [15](#), [17](#)

## I

Ignore Words [22](#)  
Insert Line [16](#)

## L

Limitations  
z/OS [10](#)

Local Population  
YLP [25](#)

## N

new rule [22](#)

## P

Population rule-set [15](#)  
Population rules [11](#)

## R

Rule Interrogation [15](#), [17](#)

## S

Selecting a System [11](#)  
SSA-NAME3  
Connecting [11](#)  
SSA-NAME3 CJK-SUPPORT [11](#)  
SSA-NAME3 Edit Rule Wizard [10](#)  
SSA-NAME3 Server [15](#)  
Standard Keys [10](#)  
Standard Population  
YSP [25](#)

## U

Unicode [17](#)