

Oracle® Enterprise Data Quality for Product Data

AutoBuild リファレンス・ガイド

リリース 5.6.2

部品番号 : B66257-01

2011 年 11 月

Oracle Enterprise Data Quality for Product Data AutoBuild リファレンス・ガイド, リリース 5.6.2

部品番号 : B66257-01

Copyright © 2001, 2011 Oracle and/or its affiliates. All rights reserved.

Primary Author: Lorna Vallad

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, the following notice is applicable:

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation shall be subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License (December 2007). Oracle America, Inc., 500 Oracle Parkway, Redwood City, CA 94065.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

This software or hardware and documentation may provide access to or information on content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services.

目次

はじめに	iii
ドキュメントのアクセシビリティ	iv
関連ドキュメント	iv
表記規則	v

1 概要

AutoBuild とスマート用語集	1-3
デフォルトのスマート用語集	1-3
生成されるデータ・レンズの基礎となるスマート用語集	1-3
スマート用語集にマージされる品目定義モデル	1-3
品目定義モデルのフレーズ構造からスマート用語集のフレーズ構造への置換	1-4
AutoBuild を使用するための前提条件	1-4
データの前提条件	1-4
構造化品目情報の特定	1-4
構造化品目コンテンツのソース	1-5
構造化情報の量	1-5
サポートされる構造化品目情報の形式	1-5
サポートされているカテゴリ形式	1-6
サポートされている属性形式	1-6

2 AutoBuild のユーザー・インタフェース

AutoBuild のユーザー・インタフェース	2-2
Step 1: Identify Category Information	2-3
「Category Source Information」 セクション	2-3
「Category/Attribute Data」 表	2-6
Category Details	2-6
Step 2: Identify Attribute Information	2-8
「Attribute Information」 セクション	2-8
「Required Attributes」 ボックス	2-8
「Scoring Attributes」 ボックス	2-12
「Optional Attributes」 ボックス	2-15
「Category/Attribute Data」 表	2-16
Attribute Details	2-16
「Attribute Values」 ボックス	2-16
「Standardizations」 ボックス	2-17

Step 3: Identify Sample Data Information	2-18
「Select Sample Data Source」 セクション	2-18
「Category/Attribute Data」 表	2-20
サンプル・データ表	2-20
Step 4: Define Data Lens Output	2-21
Outputs ボックス	2-21
AutoBuild 品目定義エクスプローラ	2-22
Output Details	2-23
AutoBuild レポート	2-25
「AutoLearning Metrics」 レポート	2-25
Semantic Knowledge Structure	2-26
AutoBuild Details Report	2-27

3 AutoBuild のプロセス

AutoBuild プロセスの使用	3-2
カテゴリ情報の特定	3-3
「List of Category Columns」 の選択	3-3
「Category Column Layout」 の選択	3-4
属性情報の特定	3-5
必須属性情報の選択	3-5
スコアリング属性情報の選択	3-6
サンプル・データ情報の特定	3-7
サンプル・データ・ソースの選択	3-7
データ・レンズ出力の定義	3-8
新規データ・レンズの完成	3-13
データ・ペアのデータ・レンズの例	3-14
その他の考慮事項	3-16

A クライアント・ソフトウェアのインストール

はじめに

このリファレンス・ガイドでは、Oracle Enterprise Data Quality for Product Data AutoBuild アプリケーションの機能について説明します。このドキュメントは、次のような構成になっています。

- 第1章および第2章では、アプリケーションの基本機能について説明します。
- 第3章では、例を使用して AutoBuild プロセスについて説明します。

このマニュアルを使用する前に、Oracle Enterprise Data Quality for Product Data の次のドキュメントをお読みください。

- Oracle Enterprise Data Quality for Product Data Knowledge Studio リファレンス・ガイド
提供されているすべての機能を理解するには、iv ページの「[関連ドキュメント](#)」に示されている Oracle Enterprise Data Quality for Product Data のドキュメントとともに、このリファレンス・ガイドを使用する必要があります。

Oracle Enterprise Data Quality for Product Data クライアント・ソフトウェアをコンピュータにインストールしておく必要があります。詳細は、[付録 A](#)、「[クライアント・ソフトウェアのインストール](#)」を参照してください。

さらに、Oracle Enterprise Data Quality for Product Data AutoBuild 製品もシステムにインストールしておく必要があります。詳細およびインストール手順は、『Oracle Enterprise Data Quality for Product Data Services for Excel リファレンス・ガイド』を参照してください。

ドキュメントのアクセシビリティ

アクセシビリティに対するオラクル社の取り組みについては、Oracle アクセシビリティ・プログラムの Web サイト (<http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc>) を参照してください。

Oracle サポートへのアクセス

Oracle のお客様は、My Oracle Support にアクセスして電子サポートを受けることができます。詳細は、<http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info> または <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs> (聴覚障害者向け) を参照してください。

関連ドキュメント

詳細については、ドキュメント・セット内の次のドキュメントを参照してください。

- Oracle Enterprise Data Quality for Product Data Oracle DataLens Server インストールेशन・ガイドには、Oracle DataLens Server のインストール手順について詳細に記載されています。
- Oracle Enterprise Data Quality for Product Data Oracle DataLens Server 管理者ガイドには、Oracle DataLens Server のインストールおよび管理について記載されています。
- Oracle Enterprise Data Quality for Product Data COM インタフェース・ガイドには、Oracle DataLens Server の COM API のインストールおよび使用について記載されています。
- Oracle Enterprise Data Quality for Product Data Java インタフェース・ガイドには、Oracle DataLens Server の Java API のインストールおよび使用について記載されています。
- Oracle Enterprise Data Quality for Product Data Application Studio リファレンス・ガイドには、Data Service Applications (DSA) の作成および保守について記載されています。
- Oracle Enterprise Data Quality for Product Data Knowledge Studio リファレンス・ガイドには、データ・レンズの作成および保守について記載されています。
- Oracle Enterprise Data Quality for Product Data Governance Studio リファレンス・ガイドには、Data Service Applications (DSA) の実行と出力データの分析に関する情報が記載されています。
- Oracle Enterprise Data Quality for Product Data 用語集には、Enterprise DQ for Product でよく使用される技術用語について記載されています。
- Oracle Enterprise Data Quality for Product Data Services for Excel リファレンス・ガイドには、Microsoft Excel ワークシートに含まれるデータに基づいて DSA を作成する方法について記載されています。
- Oracle Enterprise Data Quality for Product Data Task Manager リファレンス・ガイドには、Task Manager または Governance Studio のアプリケーションで作成されたタスクの管理について記載されています。

このドキュメントおよびすべての関連ドキュメントの最新版は、次に示す Oracle Enterprise Data Quality for Product Data ドキュメントの Web サイトから入手できます。

http://download.oracle.com/docs/cd/E20593_01/index.htm

表記規則

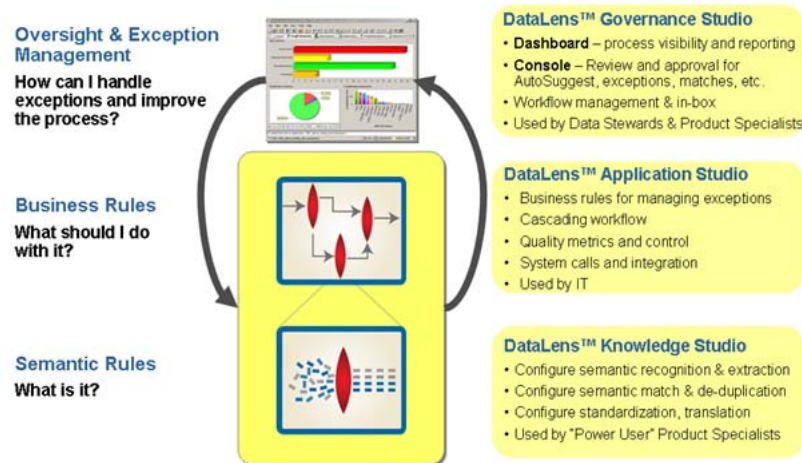
このドキュメントでは、次のテキスト表記規則を使用しています。

表記規則	意味
太字	太字は、アクションに関連するグラフィカル・ユーザー・インタフェース要素、またはテキストや用語集に定義されている用語を示します。
イタリック	イタリック体は、書籍のタイトル、強調、または特定の値を指定するブレースホルダ変数を示します。
等幅	等幅フォントは、段落内のコマンド、URL、サンプル・コード、画面に表示されるテキスト、入力するテキスト、またはファイルやディレクトリやパスの名前を示しています。
等幅	太字の等幅フォントは、入力するコマンドまたはテキストを示しています。

概要

Oracle DataLens Server は、様々なソースおよびシステムからの製品データを標準化、照合、補完および修正する、業界最先端の DataLens™ テクノロジを基に構築されています。DataLens の中核テクノロジには、製品データに特有の変動性の高さに対応するために、一から設計された特許取得済セマンティック・テクノロジが使用されています。

Oracle Enterprise Data Quality for Product Data (以前の Oracle Product Data Quality) では、DataLens の中核テクノロジ・モジュールである Governance Studio、Knowledge Studio および Application Studio が使用されます。次の図は、これらのモジュールのプロセス・フローを示しています。

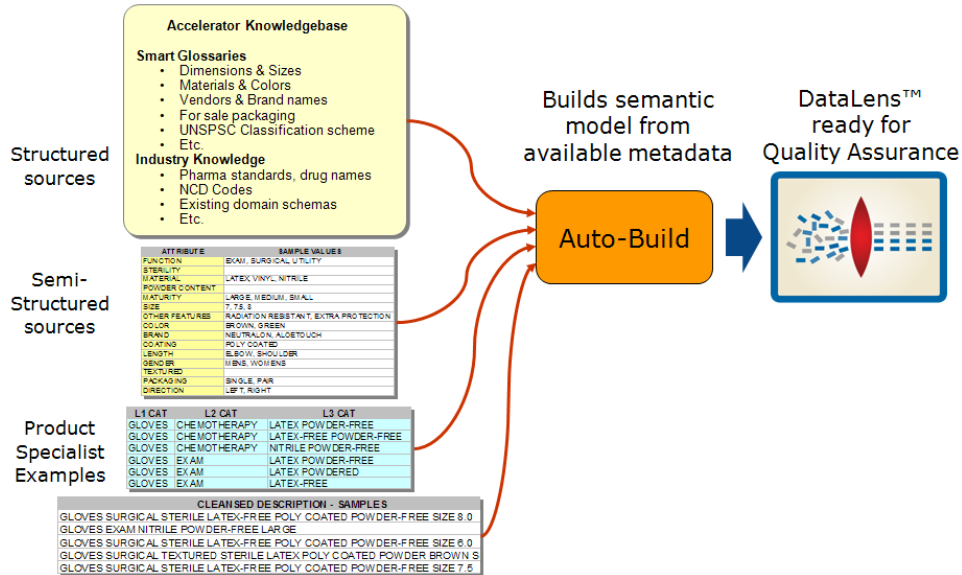


AutoBuild アプリケーションは、追加インストールする（アドオン）製品である Services for Excel に含まれており、カスタム・ツールバーとして Excel に追加されます。この製品の詳細は、『Oracle Enterprise Data Quality for Product Data Services for Excel リファレンス・ガイド』を参照してください。

AutoBuild では、Enterprise DQ for Product (EDQP) スマート用語集を使用して、既存の製品情報を利用した企業コンテンツ固有の初期データ・レンズをすばやく作成できます。たとえば、ペン、鉛筆、マジック・マーカーおよび蛍光ペンが在庫に含まれていることがすでにわかっている場合は、これらの各製品に関する情報を含む構造化コンテンツの例が会社にすでに存在する可能性があります。AutoBuild アプリケーションを使用すれば、このデータ・ナレッジを Excel ワークシートからデータ・レンズに変換できるので、労力とコストを大幅に節約できます。

AutoBuild は、製品データの例を調べて初期データ・レンズを作成します。十分な情報があれば、AutoBuild アプリケーションで次のことを行えます。

- 必須属性、スコアリング属性およびオプション属性が付いた完全な品目定義階層の作成
- 品目およびフレーズを認識するための強力なルールの作成
- 分類ルールの初期セットの提供



AutoBuild アプリケーションでは、次の作業に必要な労力を大幅に削減できます。

- 初期データ・レンズの作成
- ビジネス・ドメインに対する初期データ・レンズの調整
- データ・レンズを利用して、様々な社内外のデータ・ソース間で企業データ要件を自動的に満たす作業

AutoBuild には、最初から最後まで手順をわかりやすくガイドするグラフィカル・ウィザードが用意されています。

AutoBuild とスマート用語集

Enterprise DQ for Product スマート用語集は、広範なデータ・ドメインに適用できるように設計されたデータ・レンズです。製品とともに提供されるスマート用語集は、材料、色、単位など、一般に適用可能な項目を認識するために使用します。各スマート用語集は、既存のレンズにインポートするように設計されていますが、新規レンズを作成する際のベースとして使用することもできます。スマート用語集を組み合わせれば、データのドメインに対する認識の組合せを簡単に指定できます。スマート用語集は、Knowledge Studio で作成および編集できます。スマート用語集の詳細は、『Oracle Enterprise Data Quality for Product Data Knowledge Studio リファレンス・ガイド』を参照してください。

構造化された品目コンテンツから新規データ・レンズを AutoBuild で作成する際には、スマート用語集の組合せをベースに使用してデータ・レンズが生成されます。

デフォルトのスマート用語集

AutoBuild で使用されるデフォルトのスマート用語集は DLS_Import_Template です。このスマート用語集は、事前パッケージ済の他のスマート用語集とともにサーバーにインストールされます。このスマート用語集は、単位、数量、色、材料、仕上げ、製品パッケージの複合的な用語集です。

AutoBuild を使用してデータ・レンズを作成するたびに、自動作成プロセスで使用するスマート用語集を選択しなおすことができます。AutoBuild で使用するドメイン・データに最もよく該当するスマート用語集の独自の組合せを作成および構成することもできます。AutoBuild でデータ・レンズを作成する際に適用できるのは1つのスマート用語集のみですが、レンズを自動作成してオープンした後に追加のスマート用語集をインポートできます。

データ・レンズを生成するために AutoBuild で使用されるスマート用語集は、データ・レンズ・テンプレートとも呼ばれます。

注意： デフォルトのスマート用語集 DLS_Import_Template は、自動的に使用可能になります。データ・レンズを生成するテンプレートとして他のスマート用語集を使用する場合は、Oracle DataLens Server からチェック・アウトする必要があります。スマート用語集の詳細は、『Oracle Enterprise Data Quality for Product Data Knowledge Studio リファレンス・ガイド』を参照してください。

生成されるデータ・レンズの基礎となるスマート用語集

生成されるデータ・レンズの初期設定はすべて、AutoBuild の使用するスマート用語集から提供されるため、そのスマート用語集が基礎となります。Knowledge Studio のオプションにある標準データ・レンズのオプションが、選択したスマート用語集から新規データ・レンズにコピーされます。変換タイプが、スマート用語集から、生成されたデータ・レンズにコピーされます。これには、分類タイプ、標準化タイプおよび照合タイプが含まれます。単位変換ルールも、生成されたデータ・レンズにコピーされます。生成されたデータ・レンズには、スマート用語集に定義されているすべてのフレーズおよび用語が含まれます。

スマート用語集にマージされる品目定義モデル

AutoBuild では、選択したスマート用語集から生成された品目定義、属性、および関連するフレーズと用語のすべてが、メタデータから生成された内容にマージされます。スマート用語集のフレーズおよび用語が、指定した構造化品目情報から生成されたフレーズおよび用語と重複している場合、AutoBuild ではスマート用語集のフレーズおよび用語が保持され、生成されたフレーズおよび用語は、値セットが使用されていないかぎり無視されます。このようにして、実行時にエラーの原因となる可能性がある、あいまいな用語およびフレーズがデータ・レンズに作成されないようにします。さらに、AutoBuild では、意図しないルール間の衝突を回避するために、不必要なルールが取り除かれます。

品目定義モデルのフレーズ構造からスマート用語集のフレーズ構造への置換

AutoBuild では、生成されたフレーズの構造にドリルダウンして、構造化品目情報で見つかった単位および数タイプのフレーズ生成結果を、スマート用語集の Enterprise DQ for Product 標準フレーズ構造で置き換えることができます。スマート用語集バージョンのこれらのフレーズ構造を使用する利点は、スマート用語集のフレーズ構造がすでに標準化されていることです。ほとんどの場合、単位変換ルールはすでに適用されており、スマート用語集の品質も確認されています。

AutoBuild を使用するための前提条件

AutoBuild アプリケーションを正しく使用するためには、次の前提条件を満たす必要があります。

データの前提条件

AutoBuild アプリケーションを効率的に使用するために、会社の製品に関する次のすべてまたは一部の情報を含む Excel ワークシートの形式で、メタデータを操作する必要があります。

品目クラス

処理する品目データを表すカテゴリ。カテゴリは階層化してグループ化できますが、必須ではありません。

品目の形、適性、機能の属性

これらの属性の例としては、色、重さ、サイズ、材料、包装などが含まれます。

有効属性

データに有効な属性値の例。

構造化品目の例に、省略形ではなく、完全に表現された属性値が含まれている場合に、AutoBuild アプリケーションの効果が最も高まります。Enterprise DQ for Product では、これらの完全形の例を利用して、用語の省略形およびフレーズのバリエーションを自動的に認識できます。

構造化品目情報の特定

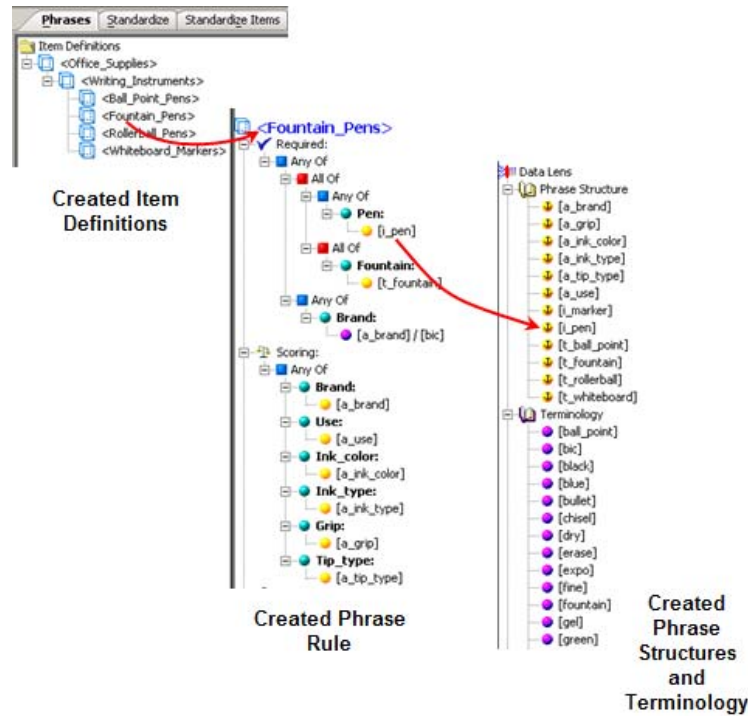
AutoBuild で使用できる構造化品目情報の主なソースは次のとおりです。

- 品目カテゴリ
- 品目名およびタイプ
- 品目ブランド情報
- 品目属性
- 属性値

サンプル・データ・ファイルの構造化品目情報の前述の各タイプの例を、次の図に示します。

SKU	Category 1	Category 2	Category 3	Name	Type	Brand	Use	Ink Color	Ink Type	Grip	Tip Type
5671231	Office Supplies	Writing Instruments	Rollerball Pens	Pen	Rollerball	Monte Blanc	Use	Blue	Gel	Standard	Bullet
5671255	Office Supplies	Writing Instruments	Fountain Pens	Pen	Fountain	Bic	Use	Black		Rubber	Fine
5678220	Office Supplies	Writing Instruments	Ball Point Pens	Pen	Ball Point	Bic	Permanent	Red		Soft	Bullet
2259871	Office Supplies	Writing Instruments	Whiteboard Markers	Marker	Whiteboard	Expo	Dry Erase	Green	Low Odor	Standard	Chisel
9383738	Office Supplies	Writing Instruments	Whiteboard Markers	Marker	Whiteboard	Expo	Dry Erase	Orange		Wide	Medium

前述の例の情報構造を前提とすると、次のようなデータ・レンズを AutoBuild で生成できます。



構造化品目コンテンツのソース

構造化品目情報のソースとして、次のものを使用できます。

- 既存の電子製品カタログ情報
- 品目マスター情報
- 製品情報管理（PIM）データ
- カテゴリ管理または製品マーケティングのワークシート
- 既存の E-Commerce サイト情報
- 現在の製品カテゴリを定義するために使用した構造化サンプル

構造化情報の量

AutoBuild では、使用可能な構造化情報を量に関係なく使用できます。カテゴリ・リストのみがある場合、AutoBuild では、指定されたカタログに準拠する一連の品目定義を含む基本データ・レンズが作成されます。カテゴリに属性名が関連付けられている場合、AutoBuild では、レンズの品目定義に属性を追加できます。属性に属性値が関連付けられている場合、AutoBuild では、フレーズおよび用語のルールを作成して、属性情報を認識し、品目定義内の正しい属性にフレーズを自動的に関連付けることができます。AutoBuild は、指定された構造化製品情報を最善の方法で利用します。

カテゴリごとにいくつかの構造化品目の例を指定すると、AutoBuild を最も効果的に使用できます。これらの品目例は、企業データで取得する必要がある製品カテゴリおよび属性を表すものにしてください。

サポートされる構造化品目情報の形式

製品情報のエクスポートは、システムとデータベース・スキーマで大きく異なる場合があります。そのため、AutoBuild は、広範な製品情報のエクスポート形式をサポートするように設計されています。

サポートされているカテゴリ形式

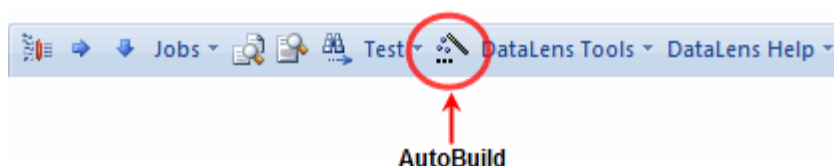
- 複数列のカテゴリ名（各カテゴリ列が分類階層の1つのレベルを表します）。
- カテゴリ列がペアでグループ化された複数列のカテゴリ・コードと名前のペア。ペアの最初の列はカテゴリ・コードで、ペアの2番目の列はカテゴリ名です。各列のペアは、分類階層の1つのレベルを表します。
- 単一系列のカテゴリ名（単一レベルの分類階層）。
- 文字で区切られた単一系列のカテゴリ名（各カテゴリ名が文字列で区切られた複数レベルの分類階層）。
- 単一系列の UNSPSC カテゴリ・コード。

サポートされている属性形式

- 属性名がカテゴリ情報と同じ行にリストされた形式。
- 属性名と属性値のペアがカテゴリ情報と同じ行にリストされた形式。3番目の列を使用して値セットを識別すれば、メタデータ・ワークシートにある値のみを使用できます。
- 属性名、値、単位（UOM）の3つ1組がカテゴリ情報と同じ行にリストされた形式。
- 属性がカテゴリ情報と同じファイルにあり、属性名が列ヘッダーとしてリストされた形式。
- 属性がカテゴリ情報とは別のファイルにあり、属性名が列ヘッダーとしてリストされた形式。
- 単一のワークシートに複数のカテゴリがあり、各カテゴリが個別の一連の行として、別のカテゴリ・グループと空白行で分けられている形式。属性がカテゴリ情報と同じファイルにあり、属性名は列ヘッダーとしてリストされます。
- カテゴリと属性情報を、同じワークシート内の複数の行に繰り返せる形式。さらに、ワークブック内の複数のワークシートにカテゴリと属性情報をリストできます。

AutoBuild のユーザー・インタフェース

AutoBuild の機能には、次に示す Services for Excel のツールバーの「AutoBuild」ボタンを使用してアクセスします。



ヒント： メニュー項目、ボタン、タブ、アイコンまたは同様のコンテンツにマウス・ポインタを合わせるとツールチップが表示されます。

ヒント： [Shift] キーを押しながら、「AutoBuild」ボタンをクリックすると、AutoBuild ユーザー・インタフェース・ウィザードのステップ 4 に直接進みます。詳細は、2-21 ページの「[Step 4: Define Data Lens Output](#)」を参照してください。

このツールバーが Excel に表示されない場合は、Services for Excel がインストールされていない可能性があります。インストールの手順と詳細は、『Oracle Enterprise Data Quality for Product Data Services for Excel リファレンス・ガイド』を参照するか、コンサルタントに問い合わせてください。

「AutoBuild」ボタンをクリックすると、次のメッセージが表示される場合があります。

```
No data lenses were found.  
Please make sure the Knowledge Studio is installed and check out the Smart Glossaries.
```

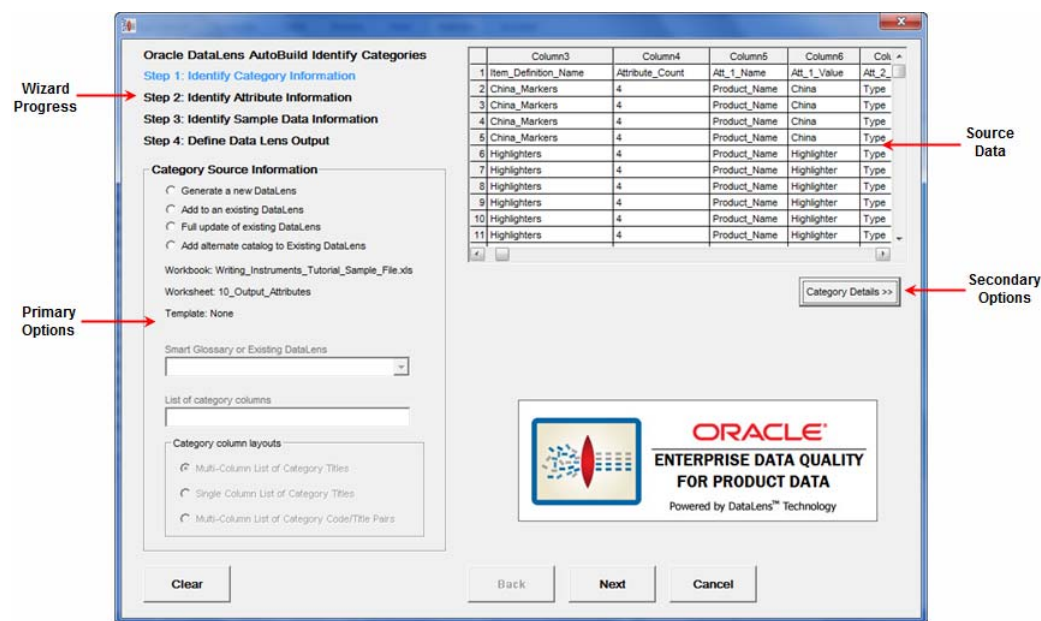
これは、Services for Excel はインストールされているが、EDQP クライアント・ソフトウェアがインストールされていないことを示しています。インストール手順については、A-1 ページの「[クライアント・ソフトウェアのインストール](#)」を参照してください。

AutoBuild のユーザー・インタフェース

AutoBuild のユーザー・インタフェースは、4つのステップのウィザードで、初期データ・レンズの作成方法を選択できるようになっています。AutoBuild ツールバーの「AutoBuild」ボタンをクリックすると、ウィザードが起動します。

ヒント： AutoBuild ウィザードを起動する前に、オープン中のすべての Excel ワークブックを保存しておくことをお勧めします。

AutoBuild ウィザードの最初のステップは「Identify Category Information」です。



注意： AutoBuild ウィザードの使用中は、アクティブな Excel ワークブックにアクセスできません。

ウィザードの各ステップには、次の4つの領域が含まれています。

ウィザードの進捗

ウィザードのアクティブ・ステップのタイトルが青色で表示されます。

ソース・データ

アクティブな Excel ワークシートのデータが表示されます。Excel の場合と同様に、列をサイズ変更でき、矢印で上下にスクロールすれば、さらにデータを表示できます。選択すると列の色が変わります。

主要オプション

この領域のコントロールでは、このステップに必要な（主要）オプションを選択できます。

追加オプション

この領域のボタンを押すと、このステップで選択できる追加オプションを有効にできます。変更しない場合、またはこの領域を有効にしない場合は、デフォルトの追加オプションが選択されます。

各ステップには、次のボタンが表示され、必要に応じて有効になります。

Clear

ウィザードの選択がリセットされ、アクティブな Excel ワークブックを使用して新規レンズ作成プロセスを起動できるようになります。

Back

ウィザードを前のステップに戻して、選択項目を確認または変更できます。

Next

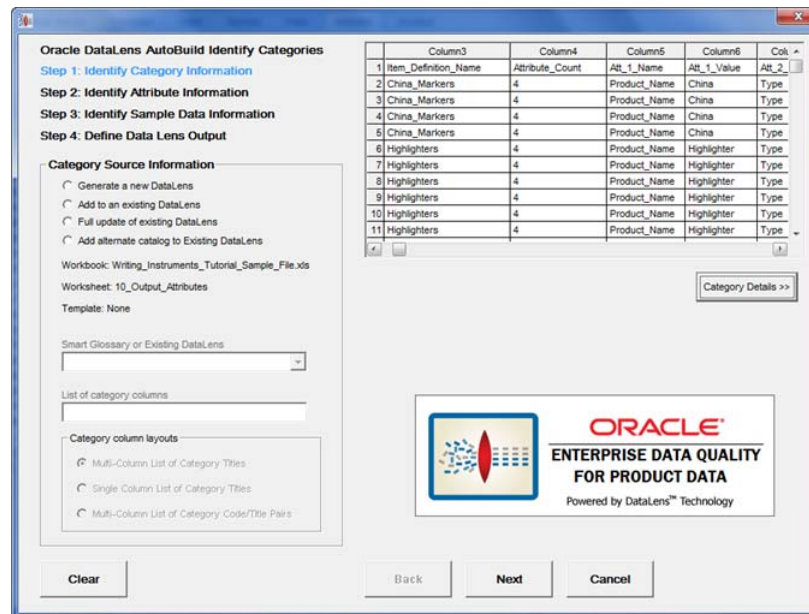
プロセスの次のステップにウィザードを進めます。

Cancel

AutoBuild ウィザードを終了します。選択項目は保存されません。

Step 1: Identify Category Information

AutoBuild ウィザードの最初のステップでは、カテゴリ情報を識別し、カテゴリ情報の使用方法を記述します。ワークシートに対して初めて AutoBuild ウィザードをオープンすると、アクティブな Excel ワークシートがデータソースとしてデフォルトで選択されます。

**「Category Source Information」 セクション**

このセクションのコントロールは次のように使用します。

データ・レンズ生成オプション

これらのオプションは、「Smart Glossary or Existing DataLens」リストの選択項目と連係して機能し、このリストに基づいてアクティブになります。AutoBuild でデータ・レンズを作成する方法を、次のオプションから選択します。

- **Generate a new DataLens**

新規データ・レンズを作成します。これは、「DLS_Import_Template」データ・レンズを選択するとアクティブになります。

「DLS_Import_Template」以外のデータ・レンズを「Smart Glossary or Existing DataLens」リストで選択すると、次の3つのオプションがアクティブになります。

- **Add to an existing DataLens**

新しい属性および品目定義が既存のデータ・レンズに追加されます。

- **Full update of existing DataLens**

新規属性および品目定義が既存のデータ・レンズに追加され、既存の属性および品目定義は更新されます。このオプションを使用して、分類スキーマを更新することもできます。

■ Add alternate catalog to Existing DataLens

選択した既存のデータ・レンズに、ユーザー定義の新規分類スキーマが追加されます。この新規代替カタログは、単に新規分類スキーマとしてデータ・レンズに追加することも、追加と同時に分類ルールが自動的に作成されるようにすることもできます。次に詳細を説明します。

データ・レンズへの代替カタログの追加（分類ルールを作成しない場合）

分類スキーマをデータ・レンズに追加できます。代替カタログは、次の形式および列名にする必要があります。

	A	B	C	D	E	F
1	catalog	parent_category_id	parent_category_desc	category_id	category_name	
2	Purchasing Catalog			2404	Electromechanical Components	
3	Purchasing Catalog	2404	Electromechanical Components	2405	Audio Devices	
4	Purchasing Catalog	2404	Electromechanical Components	2406	Circuit Breakers	
5	Purchasing Catalog	2425	IC's	2372	Memory	
6	Purchasing Catalog	2404	Electromechanical Components	2407	Fans	
7	Purchasing Catalog	2404	Electromechanical Components	2408	Industrial Controls & Sensors	
8	Purchasing Catalog	2404	Electromechanical Components	2409	Knobs & Dials	
9	Purchasing Catalog	2404	Electromechanical Components	2410	Power	
10	Purchasing Catalog	2410	Power	2411	Batteries	
11	Purchasing Catalog	2410	Power	2412	Battery Holders & Snaps	

このカタログをデータ・レンズに追加した後、（品目定義またはルールのどちらか、または両方を使用して）データ・レンズ内に分類ルールを手動で作成できます。

注意： この機能をさらに一般的に使用して、データ・レンズに追加する分類スキーマを作成することもできます。AutoBuild では、分類スキーマの形式が正しいかぎり、データ・レンズで使用できる分類スキーマが作成されます。

データ・レンズへの代替カタログの追加および分類ルールの作成

代替カタログの追加と同時に分類ルールを作成できます。この機能を実行すると、既存のデータ・レンズの品目定義を新規代替カタログ（分類スキーマ）のカテゴリにマップする分類ルールが作成されます。これを行うには、次の要件を満たす必要があります。

- 代替カタログ・スプレッドシートには、2-4 ページの「データ・レンズへの代替カタログの追加（分類ルールを作成しない場合）」に説明されている列が含まれている必要があります。
- 既存のデータ・レンズの品目定義には、値が指定された説明フィールドが必要です。
- 代替カタログ・スプレッドシートには、「mapped_primary_categories」列が含まれている必要があります。
- 代替カタログが品目定義にマッピングされている場合は、「mapped_primary_categories」列に、そのカテゴリに分類する品目定義の説明が含まれている必要があります。

これらの要件の例を次に示します。

	A	B	C	D	E	F
1	catalog	parent_category_id	parent_category_desc	category_id	category_name	mapped_primary_categories
2	Purchasing Catalog			2404	Electromechanical Components	
3	Purchasing Catalog	2404	Electromechanical Components	2405	Audio Devices	Digital Audio Devices
4	Purchasing Catalog	2404	Electromechanical Components	2406	Circuit Breakers	Breaker Panel
5	Purchasing Catalog	2425	IC's	2372	Memory	IC-Memory
6	Purchasing Catalog	2404	Electromechanical Components	2407	Fans	
7	Purchasing Catalog	2404	Electromechanical Components	2408	Industrial Controls & Sensors	
8	Purchasing Catalog	2404	Electromechanical Components	2409	Knobs & Dials	

注意： 品目定義に対するカテゴリのマッピングは品目定義の説明を使用して実行されるので、(データ・レンズ内の) 品目定義の説明フィールドに値を設定する必要があり、「mapped_primary_catgories」列の値は (品目定義の名前ではなく) 品目定義の説明を参照している必要があります。

Excel の情報

AutoBuild には、現在の Excel ワークブック、ワークシートおよびテンプレートの情報が表示されます。

Smart Glossary or Existing Data Lens List

データ・レンズの生成に使用するスマート用語集またはデータ・レンズを選択します。ドメイン・データに最もよく該当するスマート用語集の独自の組合せを作成して使用できます。

注意： データ・レンズに対する「DLS_」という接頭辞は、スマート用語集を示すために予約されているので、この命名規則を使用してデータ・レンズを選択するとスマート用語集と見なされます。たとえば、デフォルトのスマート用語集は DLS_Import_Template です。

注意： データ・レンズ名が正しく表示されない場合 (2 バイト文字の場合に発生することがあります)、Windows の地域と言語のオプションが正しく設定されていることを確認してください。詳細は、Windows のヘルプを参照してください。

「List of category columns」フィールド

このフィールドを使用して、データ・レンズの品目定義階層に使用する列を指定します。たとえば、データ・レンズの品目定義は、品目マスターの品目カタログ・カテゴリに相当します。列の順序は左から右で、親カテゴリが左、子カテゴリが右にあると仮定されています。数字、ダッシュおよびカンマの任意の組合せを使用して、列の範囲、または範囲と単一の列の組合せを指定できます。

[Tab] キーを押すか、別のユーザー・インタフェース・オプションを選択して、カテゴリリストの入力を確認する必要があります。リストを確認すると、ソース・データ表が更新されて、選択したカテゴリ列が緑色で表示されます。

「Category column layout」オプション

次の 3 つのカテゴリ列レイアウトから選択できます。

- **Multi-Column List of Category Titles**

この列レイアウトでは、カテゴリ階層に複数のレベルがあり、各列が階層の 1 つのレベルを表していると見なされます。AutoBuild では、カテゴリ階層が左から右へと表され、左の最初の列がカテゴリ階層の最高レベル、右の最後の列が階層の最低レベルを表すと見なされます。「List of category columns」フィールドに複数の列が含まれる場合は、このオプションがアクティブになり、デフォルトで選択されます。

- **Single Column List of Category Titles**

1 つのカテゴリ列のみをリストする場合は、このレイアウトがアクティブになって選択されます。AutoBuild ウィザードを最初にオープンすると、これがデフォルトになります。「List of category columns」フィールドで 1 つの列を選択すると、このオプションのみがアクティブになります。

- **Multi-Column List of Category Code/Title Pairs**

このレイアウトは「Multi-Column List of Category Titles」に似ていますが、各カテゴリが列のペアとして表されます。最初の列にはカテゴリ・コード、2 番目の列にはカテゴリ名が含まれています。リストに複数の列が含まれる場合は、このオプションがアクティブになります。

「Category/Attribute Data」表

Category/Attribute Data 表のセクションは、次のように表示されます。

Column1	Column2	Column
1 Id	Description	Item_Definition
2 WI12	China Mrkr Green Ink 12 / Box	China_Markers
3 WI13	China Markers red ink	China_Markers
4 WI14	China Markers White ink	China_Markers
5 WI15	China Mrk yellow ink	China_Markers
6 WI16	Chisel Point Highlighter Turquoise ink	Highlighters
7 WI31	Highlighter Chisel Tip Fluorescent Purple	Highlighters
8 WI32	Highlighter Chisel Tip Fluorescent Yellow	Highlighters
9 WI33	Highlighter Chisel Tip Orange	Highlighters
10 WI34	Highlighter Chisel Tip Pink	Highlighters
11 WI35	Highlighter Chisel Tip Yellow	Highlighters

Category Details <<

データ表

データ表には、構造化された Excel ワークシート・データソースの最初の 30 行と 80 列が表示されます。「List of category columns」フィールドで列が識別されると、前の例と同様に、緑色で表示されます。

Category Details

「Category Details」ボタンをクリックすると、「Category Details」セクションがアクティブになります。使用方法を次に示します。

The screenshot shows the 'Category Details' dialog box with the following options:

- No special single column category layout requirements
- Single category column contains character separated category titles
 - Character Separator:
- Single category column contains UNSPSC category codes
 - UNSPSC Reference File: ...

Options

- List of Term Libraries: View Keep Category Name Case
- Map Characters to ASCII

Buttons: Back, Next, Cancel

「List of category columns」フィールドで選択する列を 1 列のみにしないと、「Category Details」のどのオプションもアクティブになりません。これらのオプションは次のように使用します。

「No special single column category layout requirements」オプション

単一のカテゴリ列が識別されると、このオプションが自動的に選択されます。これは、列の内容が区切られていないか、カテゴリ・コードが含まれていることを示しています。

「Single category column contains character separated category titles」オプション

このオプションは、定義されたカテゴリ列に、文字で区切られた複数のカテゴリ名があることを示しています。このオプションを選択すると、次の項目がアクティブになります。

■ 「Character Separator」フィールド

前のオプションを選択した場合は、このフィールドを使用して、カテゴリ・タイトルを区切るために使用する文字をシステムに通知します。

「Single category column contains UNSPSC category codes」オプション

このオプションは、識別された単一のカテゴリ列に UNSPSC カテゴリ・コードが含まれていることを示しています。AutoBuild は UNSPSC に対応しており、次を使用して指定した UNSPSC 参照ファイルから、完全な UNSPSC 階層を自動的に作成できます。

注意： Enterprise DQ for Product には、いくつかのバージョンの UNSPSC 分類システムが同梱されています。

■ 「UNSPSC Reference File」参照ボタン

使用する UNSPSC 参照ファイルを見つけて選択するために使用します。製品に同梱されている UNSPSC 参照ファイル・セットが含まれるローカル・ディレクトリがファイル・ブラウザで自動的にオープンします。これらのファイルのいずれかを選択するか、独自のファイルを選択できます。

ウィザードのステップ 4 の分類タイプとして UNSPSC スキーマを出力するには、このオプションを選択する必要があります。詳細は、2-23 ページの「[Output Details](#)」を参照してください。

「List of Term Libraries」リスト

作成されるデータ・レンズに適用する用語ライブラリを選択します。ライブラリに含まれる用語のバリエーションおよび完全形用語のバリエーションが、生成されるデータ・レンズに追加されます。

「View」ボタン

Internet Explorer が起動し、「List of Term Libraries」リストの選択項目に基づいて、使用可能な用語ライブラリのリストが表示されます。

Keep Category Name Case

このオプションを選択すると、選択したスマート用語集またはデータ・レンズの大文字と小文字のルールではなく、ワークシートのカテゴリ名の大文字と小文字を使用して、対応する品目定義が作成されます。このオプションを選択すると、アンダースコアを使用したカテゴリ名に正しく大文字と小文字が適用されます。たとえば、「Ink Pen」は「Ink_pen」ではなく「Ink_Pen」に変更されます。

「Map Characters to ASCII」チェック・ボックス

このオプションを選択すると、データソースの発音を区別する文字（à や ñ など）が相当する ASCII 文字にマッピングされます。発音を区別する文字を含むロマンス語が入力データである際に、標準英語の同等文字が必要な場合に便利です。

Step 2: Identify Attribute Information

AutoBuild ウィザードの 2 番目のステップでは、属性情報を識別し、属性情報の使用方法を記述します。

「Attribute Information」セクション

次の項では、「Attribute Information」セクションの各オプションについて説明します。

「Required Attributes」ボックス

これらのオプションの目的は、各品目を一意に識別するために使用できる属性を特定することです。このセクションのコントロールは次のように使用します。

Add Required Attributes

このチェック・ボックスを選択すると、このボックスのコントロールがアクティブになり、ソース・データのどの列を品目定義の必須属性にするかを選択できるようになります。

- 「Item Name Column」、 「Item Type Column」 および 「Unique Name Column」 フィールド

これらのフィールドを使用して、必須ロジック名属性（品目名、品目タイプおよび一意名）を定義するために使用する列を指定します。

属性リストを指定するには、数字、ダッシュおよびカンマの任意の組合せを使用して、単一の列、列の範囲、または範囲と単一の列の組合せを指定できます。**[Tab]** キーを押すか、ステップの別のコントロールを選択して、入力を確認します。列リストを確認すると、データ表が更新されて、品目名列が紫色で表示されます。

- 「Use names and types to create required attributes」 チェック・ボックス

このオプションを選択すると、AutoBuild によって必須ロジック属性が生成されます。必須属性ロジックでは、属性が一意であることを見きわめて、必須品目セットに属性を自動的に追加しようとします。一部のスコアリングまたは形 / 適性 / 機能属性はカテゴリに対して一意、または一連の兄弟カテゴリ（同じレベルにあり同じ親カテゴリにレポートするカテゴリ）に対して一意です。

構造化ソースに広範囲のカテゴリが含まれており、ほとんどのカテゴリに属性定義が含まれている場合に、拡張された属性ロジックを使用すると最も効果的です。このような場合は、カテゴリ階層全体で属性情報の一意性をシステムが分析できます。構造化ソース情報に2つのカテゴリのみがある場合、または属性情報がカテゴリ内にまばらにある場合は、このオプションを選択しないでください。

一意名は、製品の単一のカテゴリと一意に関連付けられていると見なされるので、特定のブランド名に関連付けられた最初のカテゴリが、そのブランド名で識別されるカテゴリとして記録されます。

注意： 必須ロジックに定義する一意名は、一意な識別子であることが必要です。ブランド名がカテゴリ全体で一意でない場合は、一意名を単に品目名として使用してください。

- このオプションを次のように選択すると、ロジックが生成されます。

「Use names and types to create required attributes」 チェック・ボックスを選択し、「Attribute Information」 ボックスのコントロールを使用して属性選択を行うと、次の例で説明されているデフォルトの「Any Of」 および 「All Of」 のノードが AutoBuild によって生成されます。

例	各品目定義に対して生成される必須ロジック
品目名、品目タイプまたはブランド名が定義されておらず、「Add Required Attributes」 が選択されている場合	必須 All Of: Item_Name: Item_Type:
品目名列のみが定義されている場合	必須 All Of: Any Of: Name_Attr1: Name_Attr2: 自分の兄弟カテゴリに対して一意の属性がある場合 ...

例	各品目定義に対して生成される必須ロジック
品目タイプ列のみが定義されている場合	<p>必須</p> <p>All Of:</p> <p>All Of:</p> <p>Type_Attr1:</p> <p>Type_Attr2:</p> <p>このカテゴリのみに対して一意の属性がある場合</p> <p>...</p>
品目名と品目タイプの両方の1つ以上の列が定義されている場合	<p>必須</p> <p>All Of:</p> <p>Any Of:</p> <p>Name_Attr1:</p> <p>Name_Attr2:</p> <p>自分の兄弟カテゴリに対して一意の属性がある場合</p> <p>...</p> <p>All Of:</p> <p>Type_Attr1:</p> <p>Type_Attr2:</p> <p>このカテゴリのみに対して一意の属性がある場合</p> <p>...</p>
1つ以上のブランド名列が定義されている場合	<p>必須</p> <p>Any Of:</p> <p>All Of:</p> <p>Item_Name:</p> <p>Item_Type:</p> <p>Any Of:</p> <p>Brand:</p> <p>[a_brand]¥[some_brand_name1]</p> <p>[a_brand]¥[some_brand_name2]</p> <p>...</p>

例	各品目定義に対して生成される必須ロジック
1 つ以上のブランド名列および品目名列と品目タイプ列が定義されている場合	<p>必須</p> <p>Any Of:</p> <p>All Of:</p> <p>Any Of:</p> <p>Name_Attr1:</p> <p>Name_Attr2:</p> <p>自分の兄弟カテゴリに対して一意の属性がある場合</p> <p>...</p> <p>All Of:</p> <p>Type_Attr1:</p> <p>Type_Attr2:</p> <p>このカテゴリのみに対して一意の属性がある場合</p> <p>...</p> <p>Any Of:</p> <p>Brand:</p> <p>[a_brand]¥[some_brand_name1]</p> <p>[a_brand]¥[some_brand_name2]</p> <p>...</p>

■ このオプションが選択されていない場合に生成されるロジック

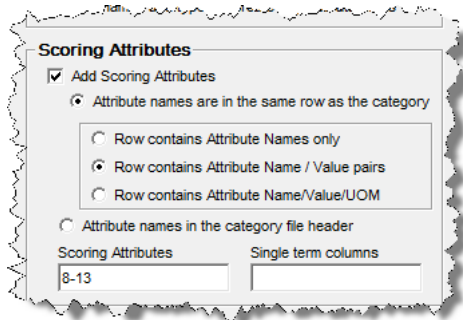
「Use names and types to create required attributes」チェック・ボックスが選択されていない場合に AutoBuild によって生成される必須ロジック属性を次の例で説明します。

例	各品目定義に対して生成される必須ロジック
品目名列または品目タイプ列が定義されていない場合	<p>必須</p> <p>All Of:</p> <p>Item_Name:</p> <p>Item_Type:</p>
品目名列のみが定義されている場合	<p>必須</p> <p>All Of:</p> <p>Item_Name:</p> <p>[i_column1]</p> <p>[i_column2]</p> <p>...</p> <p>Item_Type:</p>
品目タイプ列のみが定義されている場合	<p>必須</p> <p>All Of:</p> <p>Item_Name:</p> <p>Item_Type:</p> <p>[t_column1]</p> <p>[t_column2]</p> <p>...</p>

例	各品目定義に対して生成される必須ロジック
品目名と品目タイプの両方の 1 つ以上の列が定義されている場合	<p>必須</p> <p>All Of:</p> <p>Item_Name:</p> <p>[i_column1]</p> <p>[i_column2]</p> <p>...</p> <p>Item_Type:</p> <p>[t_column1]</p> <p>[t_column2]</p> <p>...</p>
1 つ以上のブランド名列が定義されている場合	<p>必須</p> <p>Any Of:</p> <p>All Of:</p> <p>(以前に説明した品目名および品目タイプのロジック)</p> <p>Brand:</p> <p>[a_brand]¥[some_brand_name1]</p> <p>[a_brand]¥[some_brand_name2]</p> <p>...</p>

「Scoring Attributes」ボックス

次の項では、「Scoring Attributes」ボックスの各オプションについて説明します。



「Add Scoring Attributes」チェック・ボックス

新規データ・レンズの品目定義に、スコアリング（形、適性および機能）属性を追加するには、このオプションを選択します。

「Attribute names are in the same row as the category」オプション

関連付けられているカテゴリと属性情報が同じ行にある場合は、このオプションを選択します。この場合、1つ以上の列に異なるカテゴリ・レベルが含まれており、同じワークシートに、属性名と属性値の例などが含まれる隣接する列があります。属性列のヘッダーは、通常、「Attribute Name 1」または「Attribute Value 1」などの一般的なものです。各行にある属性値情報は、次の例に示すように、その行によって定義されるカテゴリ固有の情報です。

Category 1	Category 2	Category 3	Category 4	AttrName1	AttrValue1	AttrName2	AttrValue2	AttrName3	AttrValue3
Office Supplies	Writing Instruments	Pens	Rollerball	Ink Color	Blue	Grip	Standard	Point	Fine
Office Supplies	Writing Instruments	PENS	Fountain	Ink Color	Black	Grip	Rubber	Point	Fine
Office Supplies	Writing Instruments	Pens	Ball Point	Ink Color	Red	Grip	Soft	Point	Fine

Category column
Category specific attribute names and values
Generic attribute column

次の3つのオプションは、このオプションが選択されている場合にのみアクティブになります。

■ 「Row contains Attribute names」オプション

次の例に示すように、構造化データ・ソースで属性値ではなく属性名が使用可能であることを AutoBuild に指示する場合に、このオプションを選択します。

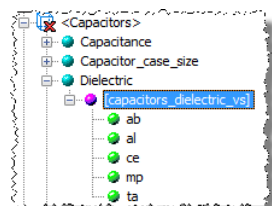
Category 1	Category 2	Category 3	Category 4	AttrName1	AttrName2	AttrName3
Office Supplies	Writing Instruments	Pens	Rollerball	Ink Color	Grip	Point
Office Supplies	Writing Instruments	PENS	Fountain	Ink Color	Grip	Point
Office Supplies	Writing Instruments	Pens	Ball Point	Ink Color	Grip	Point
Office Supplies	Writing Instruments	Markers	Whiteboard	Ink Color	Grip	Point

■ 「Row contains attribute name / value columns」オプション

次の例に示すように、属性名と属性値の両方がペアの列として、関連する属性列に含まれていることを AutoBuild に指示する場合に、このオプションを選択します。

C	D	E	F	G	H	I
catalog_group_id	catalog_group	attr_id	attr_name	attr_value	attr_uom	valueset
570	Passives					
571	Capacitors	Capacitor_SpecsDielectric	Dielectric	AB		Y
571	Capacitors	Capacitor_SpecsDielectric	Dielectric	AL		Y
571	Capacitors	Capacitor_SpecsDielectric	Dielectric	CE		Y
571	Capacitors	Capacitor_SpecsDielectric	Dielectric	MP		Y
571	Capacitors	Capacitor_SpecsDielectric	Dielectric	TA		Y
571	Capacitors	Capacitor_SpecsCAPACITANCE	CAPACITANCE	90	uF	
571	Capacitors	Capacitor_SpecsCAPACITANCE	CAPACITANCE	*00022	uF	

値セットを使用して作成される用語の名前の前に「_vs」が付き、スマート用語集の用語は存在しても無視されます。前の例の属性から、[capacitors_dielectric_vs]の用語が作成され、ab、al、ce、mpおよびtaの値セットが含まれます。



値セットが「Yes」と「No」（またはYとNのような同等のバリエーション）の場合は、「not」を使用するように、これらの値が標準化されます。たとえば、[capacitor_networks_shielded_vs]の用語の値セットが、YとNを使用して設定されていれば、データ・レンズの値セットは「shielded」と「not shielded」に標準化されます。

注意： AutoBuild を使用して、「valueset」列を含むメタデータを使用してデータ・レンズを作成する場合、「valueset」列は選択しないでください。属性名、値、単位の3つの列を選択すると、「valueset」列が AutoBuild で自動的に認識されて適切に解釈されます。

■ 「Row contains attribute name / value / UOM columns」 オプション

次の例に示すように、属性名、属性値および単位が3列のセットとしてグループ化されて、関連する属性列に含まれていることを AutoBuild に指示する場合には、このオプションを選択します。

Category 1	Category 2	Category 3	Category 4	AttrName1	AttrValue1	AttrUOM1	AttrName2	AttrValue2	AttrUOM2
Hardware	Fasteners	Screws	Wood	Material	Stainless Steel		Point Size	8 points	
Hardware	Fasteners	Bolts	Lag	Material	Stainless Steel		Shear Strength	2500 lbs	
Hardware	Fasteners	Bolts	Toggle	Material	Stainless Steel		Shear Strength	1500 lbs	
Hardware	Fasteners	Bolts	Carriage	Material	Stainless Steel		Shear Strength	3500 lbs	

ソース・データに attr_uom という名前の単位列が含まれている場合は、このオプションが自動的に選択されて、「Single term columns Scoring Attribute」 オプションが強制的に選択されます。

「Attribute names in the category file header」 オプション

構造化データ・ソースのヘッダー行（行1）に属性名がある場合には、このオプションを選択します。単一のカテゴリ、または同じ属性情報を共有する密接に関連したカテゴリ・セットがワークシートに含まれている場合には、このワークシート構造をよく使用します。たとえば、様々なタイプのペンがあり、各ペン・タイプは同じ属性を共有しているが、固有の様々なタイプのペンがあるような場合に使用します。

1つの表で3つの関連カテゴリがあるワークシートの例を次に示します。

Category 1	Category 2	Category 3	Use	Ink Color	Ink Type	Grip	Tip Type
Office Supplies	Writing Instruments	Rollerball Pens		Blue	Gel	Standard	Bullet
Office Supplies	Writing Instruments	Fountain Pens		Black		Rubber	Fine
Office Supplies	Writing Instruments	Ball Point Pens	Permanent	Red		Soft	Bullet
Office Supplies	Writing Instruments	Whiteboard Markers	Dry Erase	Green	Low Odor	Standard	Chisel
Office Supplies	Writing Instruments	Whiteboard Markers	Dry Erase	Orange		Wide	Medium

次の例に示すように、（少なくとも最初の5列について）各表のヘッダー行の属性名が同じであれば、様々なソース・データを含むいくつかの表を処理できます。

Description	Level 1	Level 2	Level 3	Product Name	Type	Ink_color	Size	Ink_Type
W104 Pen Rll Med Bk 10/Bx	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Rll	Rollerball	Ink Black		
W193 Refill Rollerball Ink Black	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Rollerball	Ink Blue		
W194 Refill Rollerball Ink Blue	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Roller Ball	Blue ink		
W103 Roller Ball Pen Refill Medium Point Blue Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Roller Ball	Blue ink		
W106 Ballpoint Pen Medium Point Black Usable barrel ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Ballpoint	Black ink		
W117 Comfort Grip Retractable Ballpoint Pen Medium Point Blue ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Retractable Ballpoint	Blue ink		
W118 Comfort Grip Retractable Ballpoint Pen Medium Point Red Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Retractable Ballpoint	Red ink		
W127 Retractable Ballpoint Pen Non-Refillable Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Retractable Ballpoint	Black Ink		
W128 Erasable Ink Ballpoint Pen Black Point Blue Barrel/Blue Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Ballpoint	Blue ink		Erasable Ink
W129 Felt Tip Pen Extra-Fine Point Translucent blue Barrel Purple in Writing_Instruments	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Felt Tip	Purple Ink		
W130 Felt Tip Pen Extra-Fine Point Translucent black Barrel Red Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Felt Tip	Red Ink		
W179 Pen Extra-Fine Point Blue Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Blue Ink	Green Ink		
W180 Pen Extra-Fine Point Green Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Green Ink			
W188 Porous Pen Extra-Fine Point Red Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Porous	Red Ink		
W191 Porous Pen Extra-Fine Tip White Barrel/Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Porous	Black Ink		
W195 Ballpoint Pen Retractable Fine Point Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Ballpoint	Black Ink		
W196 Retractable Ballpoint Pen Fine Point Clear Barrel/Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Retractable Ballpoint	Black Ink		
W197 Retractable Ballpoint Pen Medium Point Black Barrel/ black in Writing_Instruments	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Retractable Ballpoint	Black Ink		
W198 Retractable Ballpoint Pen Refillable Medium Point Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Retractable Ballpoint	Black Ink		
W199 Ballpoint Pen Rubber Grip Fine Point Retractable Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Retractable Ballpoint	Black Ink		
W100 Retractable Ballpoint Pen Rubber Grip Fine Point Red Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Retractable Ballpoint	Red Ink		
W101 Roller Ball Pen Gel ink Refillable Medium Blue	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Roller Ball	Blue ink		Gel ink
W102 Roller Ball Pen Refillable Medium Red gel ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Roller Ball	Red ink		Gel ink
W107 Pen - Ballpoint - Blue Ink, Medium Point	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Pen	Ballpoint	Blue Ink		
Description	Level 1	Level 2	Level 3	Product Name	Type	Ink_color	For Sale	Packaging
W112 China Mkr Green Ink 12/ Box	Writing_instruments	Markers	China_markers	China	Marker	Green Ink		12
W113 China Markers red ink	Writing_instruments	Markers	China_Markers	China	Markers	red ink		
W114 China Markers White ink	Writing_instruments	Markers	China_Markers	China	Markers	White ink		
W115 China Mkr yellow ink	Writing_instruments	Markers	China_Markers	China	Mkr	yellow ink		
Description	Level 1	Level 2	Level 3	Product Name	Type	Ink_color	For Sale	Packaging
W116 Chisel Point Highlighter Turquoise ink	Writing_instruments	Markers	Highlighters	Highlighter	Chisel Point	Turquoise ink		
W131 Highlighter Chisel Tip Fluorescent Purple	Writing_instruments	Markers	Highlighters	Highlighter	Chisel Tip	Fluorescent Purple		

「Scoring Attributes」フィールド

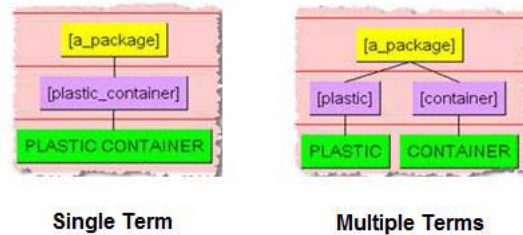
このフィールドを使用して、スコアリング属性を定義するために使用する列を指定します。属性の追加順序は、最小列番号から最大列番号の順になります。属性リストを指定するには、数字、ダッシュおよびカンマの任意の組合せを使用して、単一の列、列の範囲、または範囲と単一の列の組合せを指定できます。[Tab] キーを押すか、別のユーザー・インタフェース・オプションを選択して、属性リストの入力を確認します。列リストを確認すると、「Category/Attribute Data」表が更新されて、選択したカテゴリ列が緑色で表示されます。

「Single term columns」フィールド

このオプションを使用すると、複数の単語を含む属性値に対して、単一の用語が強制的に作成されます。識別された属性値に対するデータ・レンズのフレーズおよび用語が AutoBuild で作成される際には、属性値が単一の用語列として識別されるので、属性値にある単語（トークン）の数に関係なく、属性を認識するために 1 つの用語が AutoBuild で作成されます。たとえば、属性値が「pens」の場合に、このオプションを設定しても効果はありません。しかし、属性値が「rollerball pens」の場合は、両方の単語を使用して「[rollerball_pens]」という単一の用語が作成されます。

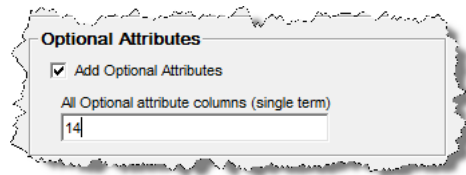
注意： 予測機能および AutoSuggest 機能は、用語が単一用語化されているときに最も効果的です。一般に、テキストの列は単一用語化されますが、単位の列は単一用語化されません。

次の例は、「PLASTIC」と「CONTAINER」という 2 つの単語が属性値である場合に、単一用語と複数用語でどのように用語が作成されるかを示しています。



「Optional Attributes」ボックス

次の項では、「Optional Attributes」ボックスの各オプションについて説明します。



「Add Optional Attributes」チェック・ボックス

オプションの属性を追加して、「Add Optional Attributes columns (single term)」チェック・ボックスを有効にするには、このオプションを選択します。

「Add Optional Attributes columns (single term)」チェック・ボックス

「Manufacturer」、「Brand_or_Model_Name」、「Model_or_Part_Number」、「Units_per_Package」、「Additional_Information」、「URL」など、すべての品目に共通の主要ビジネス属性をデータ・レンズに追加するには、このオプションを選択します。オプション属性は、フィールドに指定したすべての属性に対して作成されます。

「Category/Attribute Data」表

「Category/Attribute Data」表のセクションは、次のように表示されます。

	Column3	Column4	Column5	Column6	Column7	Column
1	Item_Definition_Na	Attribute_	Att_1_Name	Att_1_Value	Att_2_Name	Att_2_Value
2	China_Markers	4	Product_Name	China	Type	Mrkr
3	China_Markers	4	Product_Name	China	Type	Markers
4	China_Markers	4	Product_Name	China	Type	Markers
5	China_Markers	4	Product_Name	China	Type	Mrk
6	Highlighters	4	Product_Name	Highlighter	Type	Chisel Poin
7	Highlighters	4	Product_Name	Highlighter	Type	Chisel Tip
8	Highlighters	4	Product_Name	Highlighter	Type	Chisel Tip
9	Highlighters	4	Product_Name	Highlighter	Type	Chisel Tip
10	Highlighters	4	Product_Name	Highlighter	Type	Chisel Tip
11	Highlighters	4	Product_Name	Highlighter	Type	Chisel Tip

Attribute Details >>

データ表

データ表には、構造化された Excel ワークシート・データソースの最初の 30 行と 80 列が表示されます。カテゴリおよび属性の情報を識別すると、前の例に示すように適切な列が色分けして表示されます。カテゴリ列は、次の色で表示されます。

色	意味
緑	カテゴリ
紫	品目名、品目タイプおよびブランド名
青	スコアリング属性

Attribute Details

「Attribute Details」ボタンをクリックすると、「Attribute Details」セクションがアクティブになります。使用方法を次に示します。

Attribute Details <<

Attribute Details

Attribute Values

Phrase Separator: Include attribute name in phrase for context:

Maximum Terms in a Phrase: Append as needed

Map values directly to each attribute Insert always

Attribute values are Part Number Codes Append always

Never

Standardizations

#Required Standardization Columns: Create Replace All Rules

#Scoring Standardization Columns:

#Optional Standardization Columns:

「Attribute Values」ボックス

次の項では、「Attribute Values」ボックスの各オプションについて説明します。

「Phrase Separator」フィールド

このフィールドを使用して、1つの属性値列にある複数の属性を区切るために使用する文字をリストします。このリストは、句読文字を空白で区切ったリストにする必要があります。AutoBuild では、定義された文字セパレータのいずれかで区切られた値ごとに、個別の属性値が作成されます。デフォルトでは、フレーズ区切文字リストはセミコロン (;) に設定されます。

「Maximum #Terms in a Phrase」フィールド

このフィールドを使用して、関連するフレーズに AutoBuild が追加を試みる属性値の用語の最大数を制御します。属性値の用語の数がこのフィールドの数字よりも大きい場合は、余分な属性値が無視されます。このオプションには 1 から 50 までの数字を設定する必要があります。デフォルト値は 5 です。

「Map values directly to each attribute」チェック・ボックス

属性値の定義および構造化が良好で（最適データとみなされる）、これらのデータのみを割り当てようとする場合は、このチェック・ボックスを選択します。生成または更新されるデータ・レンズでは、品目定義のスコアリング属性に対して完全なフレーズおよび用語コンテキストが関連付けられます。

「Attribute values are Part Number Codes」チェック・ボックス

このチェック・ボックスを使用して、属性値が部品番号であることを示します。

「Include attribute name in phrase for context」オプション

一意の属性値認識を作成するために AutoBuild によって属性値に追加される、コンテキストを制御する 3 つのオプションのいずれかを選択します。AutoBuild では、関連するフレーズ構造に属性名が追加されて、フレーズ間のあいまいさを排除して、一意に属性値を認識できるようにします。

- **Append as needed**

属性値が一意でない場合、または属性値が数字または単位である場合のみに、フレーズ構造の最後に属性名を追加するには、このオプションを選択します。

- **Append always**

すべての属性値のフレーズ構造の最後に属性名を追加するには、このオプションを選択します。

- **Insert always**

すべての属性値のフレーズ構造の最初に属性名を挿入するには、このオプションを選択します。

- **Never**

すべてのフレーズ構造に属性名が含まれないようにするには、このオプションを選択します。

「Standardizations」ボックス

次の項では、「Standardizations」ボックスの各オプションについて説明します。

「#Required/Scoring/Optional Standardization Columns」リスト

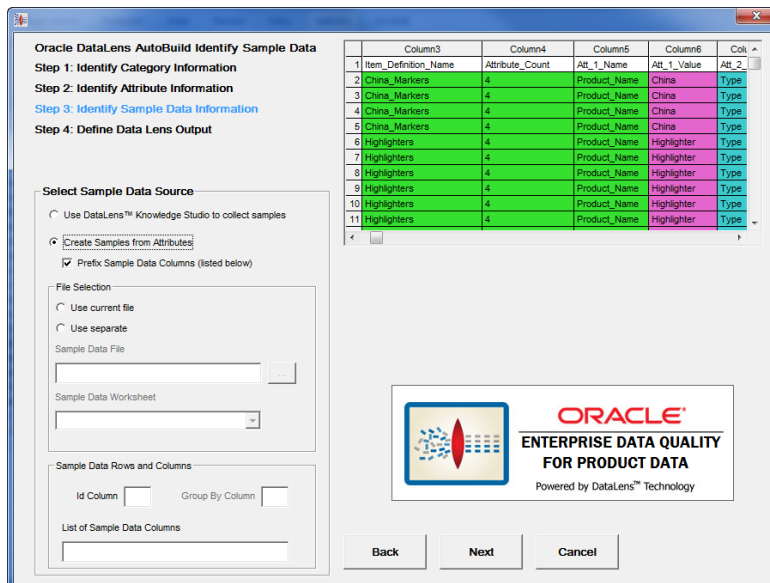
属性値のターゲット標準化フォームがメタデータ・ワークシートに含まれている場合は、この情報を使用して、自動作成されるデータ・レンズ内に標準化ルールを自動的に作成できます。これらのオプションでは、対応する標準化属性値を含む各関連属性列の右側に表示される標準化列の数を選択できます。識別される標準化列ごとに、列ヘッダーを使用して、新規レンズ内に一意の標準化タイプが作成されます。

「Create Replace All Rules」チェック・ボックス

デフォルトの標準化で、すべての属性に対して標準化ルールが作成されます。これらのルールはすべて「Replace all」のタイプになります。

Step 3: Identify Sample Data Information

AutoBuild ウィザードの3番目のステップでは、サンプル・データ情報を識別し、カテゴリ情報の使用方法を記述します。



「Select Sample Data Source」セクション

次の項では、「Select Sample Data Source」セクションの各オプションについて説明します。

「Use Knowledge Studio to collect samples」オプション

Knowledge Studio を使用してデータ・レンズ保守用のサンプル・データを収集する場合には、このオプションを選択します。AutoBuild プロセス中にサンプル・テキストを即座に使用できない場合に、この選択項目を使用します。

「Create Samples from Attributes」オプション

AutoBuild ウィザードを使用して識別された属性値からサンプル・データを直接作成する場合には、このオプションを選択します。サンプル・テキストが使用できず、既存の属性値が、Oracle DataLens Server で処理する必要があるデータを代表していると考えられる場合に、この選択項目を使用します。このオプションを選択すると、「Sample Data Rows and Columns」ボックスおよび次のチェック・ボックスがアクティブになります。

■ 「Prefix Sample Data Columns (listed below)」チェック・ボックス

サンプル構造化データの他の列にある、名前、タイプ、ブランドまたは短い説明が指定された属性値から作成されるサンプル・データに接頭辞を付ける必要がある場合には、このオプションを選択します。このチェック・ボックスを選択した場合は、「List of Sample Data Columns」フィールドを使用して、属性値を接頭辞として付ける列を定義する必要があります。

「File Selection」ボックス

このボックスには、サンプル・データが含まれるファイルやワークシートを識別するために使用するオプションが含まれています。ファイルを選択すると、サンプル・データ・ファイルの表示例が更新され、選択したオプションを反映した色で列が表示されます。サンプル構造化データのソースとして現在使用しているファイルも選択できます。

■ 「File Selection」ボックス

このボックスには、サンプル・データを選択するために使用するオプションが含まれています。

– 「Use current file」オプション

このオプションを使用すると、使用するワークシートをオープン中の Excel ワークブックから選択する「Sample Data Worksheet」リストがアクティブになります。

– 「Use separate」オプション

このオプションを使用すると、「Sample Data File」フィールドがアクティブになります。このフィールドでは、サンプル・データ・ファイルの完全なファイル・パスを入力するか、このフィールドの隣にあるボタンをクリックしてファイルを見つけます。

– 「Sample Data File」フィールド

このフィールドには、現在選択中の Excel ワークブック（または Excel 互換ファイル）への完全なファイル・パスが表示されます。これを使用して、使用する属性ディレクトリ（フォルダ）のパス名を見つけて選択するか入力します。

– 「Sample Data Worksheet」リスト

このリストを使用して、選択した Excel ワークブックで使用するワークシートを選択します。

■ 「Sample Data Rows and Columns」ボックス

このボックスには、サンプル・データ入力ファイルの列を使用して、サンプル・データ・ファイルを作成する方法を指定するために使用するオプションが含まれています。

– 「Id Column」フィールド

サンプル・データ・ソースにある一意の ID 列を識別するには、このオプションを使用します。属性値からサンプル・データを作成する際には、一意の ID が必要です。この値は、単一の列の数字にする必要があります。

– 「Group By Column」フィールド

指定した列にある一意の値によってグループ化したサンプル・ファイルを作成するために使用する列を指定するには、このオプションを使用します。AutoBuild のデフォルトでは、指定したソース・ファイルからランダムに選択されたサンプル・ファイルの標準 100 レコードのデータ・レンズが作成されます。このオプションを指定すると、グループ別ファイルも作成されます。このオプションを使用する利点は、カテゴリの専門家が、1 度に 1 つの関連品目セットのレンズを保守できることです。属性値からサンプルを作成する際にはこのオプションを使用できません。属性値からサンプルを作成する際には、指定したカテゴリ情報からグループ別サンプルが自動的に作成されます。

– 「List of Sample Data Columns」フィールド

このオプションを使用して、データ・サンプルを作成するために連結する列を指定します。指定した列のデータは、各列値の間に空白を入れて連結されます。列は列番号の小さい列から大きい列の順序で連結されます。列リストを指定するには、数字、ダッシュおよびカンマの任意の組合せを使用して、単一の列、列の範囲、または範囲と単一の列の組合せを指定できます。たとえば、5-8 または 5,6,8,9、5-8,10 のように指定できます。[Tab] キーを押すか、別のユーザー・インタフェース・オプションを選択して、列リストの入力を確認します。

「Category/Attribute Data」表

次に、「Identify Attribute Information」ダイアログの「Category/Attribute Data」表について説明します。

	Column3	Column4	Column5	Column6	
1					
2	Level 1	Level 2	Level 3	Product_Name	T
3	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
4	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
5	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
6	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
7	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
8	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
9	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
10	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
11	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B

データ表

データ表には、構造化された Excel ワークシート・データソースの最初の 30 行と 80 列が表示されます。カテゴリおよび属性の情報を識別すると、前の例に示すように適切な列が色分けして表示されます。カテゴリ列は、次の色で表示されます。

色	意味
緑	カテゴリ
紫	品目名、品目タイプおよびブランド名
青	スコアリング属性

サンプル・データ表

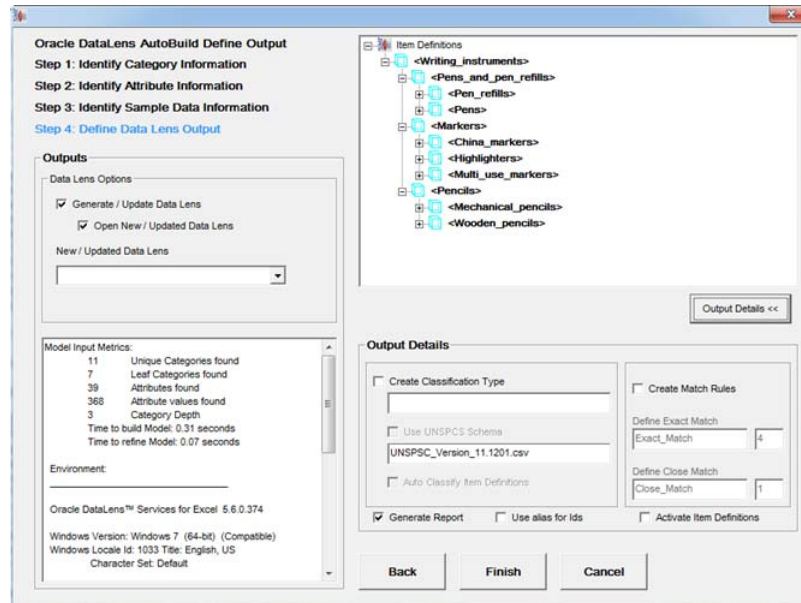
次に、「Sample data file」オプションが選択されているときに表示されるサンプル・データ表について説明します。この表の機能は、「Category/Attribute Data」表に似ています。

	Column3	Column4	Column5	Column6	Col
1					
2	Level 1	Level 2	Level 3	Product_Name	Type
3	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Ballpoint
4	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Ballpoint
5	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Ballpoint
6	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Ballpoint
7	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Ballpoint
8	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Ballpoint
9	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Ballpoint
10	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Ballpoint
11	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Rollerball

Column1	Column2	Column
Id	Description	Level 1
WI7	Ballpoint Pen Refill Fine Point Blue Ink 2 / Pack	Writing_Instrun
WI8	Ballpoint Pen Refill Med Pt Black Ink 2 / Pk	Writing_Instrun
WI9	Ballpoint Pen Refill Med Pt Black Ink 2 / Pack	Writing_Instrun

Step 4: Define Data Lens Output

AutoBuild ウィザードの最後のステップでは、データ・レンズ出力を定義します。この項では、その使用方法を説明します。



Outputs ボックス

このセクションのコントロールは次のように使用します。

「Generate / Update Data Lens」チェック・ボックス

「Finish」ボタンを選択したときにデータ・レンズが作成または更新されるようにするには、このチェック・ボックスを選択します。次のチェック・ボックスがアクティブになります。

「Open New / Update Data Lens」チェック・ボックス

データ・レンズの作成後に新規または更新済のデータ・レンズが Knowledge Studio でオープンされるようにするには、このチェック・ボックスを選択します。

「New / Updated Data Lens」フィールド

このフィールドは、新規データ・レンズに名前を付けたり、追加して更新する既存のデータ・レンズを選択するために使用します。新規データ・レンズを作成する場合は一意の名前にする必要がありますが、新規データ・レンズ名のベースとして使用するために既存のデータ・レンズを選択することもできます。

レポート・フィールド

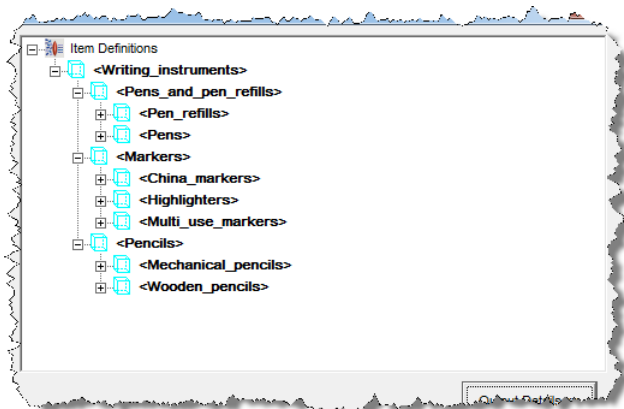
この読取り専用フィールドには、新規データ・レンズ用に生成されるモデルのメトリックを示すレポートが含まれています。AutoBuild では、モデル生成プロセスに関するすべての警告メッセージがこのフィールドにレポートされます。たとえば、カテゴリ名の重複はその一例です。

警告メッセージやエラー・メッセージがレポートされている場合は、時間を取ってメッセージを確認し、続行する前に構造化データ入力を訂正する必要があるかどうかを見極めます。

AutoBuild は、警告メッセージの有無に関係なく、「Finish」ボタンが選択されると、データ・レンズを生成します。警告メッセージは、構造化データが AutoBuild で正しく解釈されなかった可能性があることを示しています。

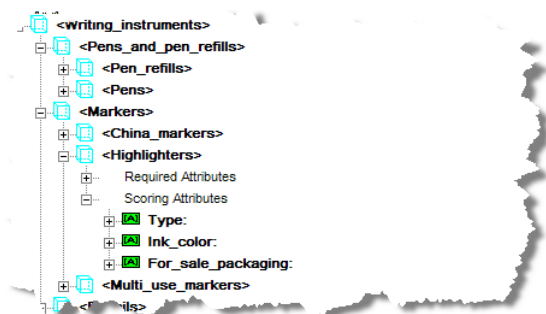
AutoBuild 品目定義エクスプローラ

品目定義エクスプローラでは、AutoBuild によって作成されたレンズ・モデルを確認できます。生成されたモデルには、指定した構造化データと選択したオプションが反映されています。



AutoBuild 品目定義エクスプローラでは、データ・レンズを生成する前に、選択項目の結果を確認できます。モデルに不整合を見つけた場合、データ・レンズが正しく作成されるように、AutoBuild ウィザードの前のステップに戻って必要な調整または訂正を加えることができます。

AutoBuild 品目定義エクスプローラには、品目定義の階層が、インデントされた一連のノードとして表示されます。青い立方体のアイコンが付いた各ノードは、品目定義または品目定義階層内のレベルを示しています。エクスプローラ内の任意のノードで、プラス・アイコンをクリックすると詳細が表示され、マイナス・アイコンをクリックすると詳細が非表示になります。



各フレーズをさらに深くドリルダウンすると、最後にフレーズの生成結果が1つ以上のグループ化された用語名のセットとして表示されます。各属性については、割り当てられたフレーズを表示できます（ある場合）。次の例に示すように、フレーズには、通常、関連付けられたレンズ・フレーズを示す黄色のボール・アイコンがあります。



Output Details

「Output Details」 ボタンをクリックすると、「Output Details」 セクションがアクティブになります。使用方法を次に示します。

「Create Classification Type」 チェック・ボックス

新規レンズに分類を追加するには、このチェック・ボックスを選択します。このチェック・ボックスを選択すると、その下のフィールドおよび「Auto Classify Item Definitions」 チェック・ボックスがアクティブになります。

- 名前フィールド

このフィールドを使用して、新規分類タイプに名前を付けます。

- 「Use UNSPSC Schema」 チェック・ボックス

2-3 ページの「[Step 1: Identify Category Information](#)」で「**Single category column contains UNSPSC category codes**」 オプションを設定した場合は、このチェック・ボックスがアクティブになります。

UNSPSC カテゴリ・コードでカテゴリが定義されている場合は、この選択項目を使用して、定義された分類タイプのスキーマ・ファイルとして、関連する UNSPSC リファレンス・スキーマを追加できます。このチェック・ボックスを選択すると、選択した UNSPSC リファレンス・スキーマ・ファイルの名前が、チェック・ボックスの下の読取り専用フィールドに表示されます。UNSPSC スキーマの構成の詳細は、2-6 ページの「[Category Details](#)」を参照してください。

- 「Auto Classify Item Definitions」 チェック・ボックス

このフィールドを使用すると、(品目定義構造と同じ) 分類タイプおよび適切な分類ルールが AutoBuild で作成されます。

「Generate Report」 チェック・ボックス

このチェック・ボックスを使用すると、2-18 ページの「[Step 3: Identify Sample Data Information](#)」に説明されているすべてのモデル・メトリック・レポートにあるすべての情報を含む AutoBuild メトリック・レポート、生成済データ・レンズの詳細、および AutoBuild セッションで使用されたオプション選択の完全なリストを生成できます。詳細は、2-25 ページの「[AutoBuild レポート](#)」を参照してください。

Use alias for Ids

カテゴリ ID および属性 ID を別名フィールドに設定する場合は、このチェック・ボックスを選択します。

これらの一意のカテゴリ ID および属性 ID は、system_id および group_id 内部フィールドに保存され、「Item Definition Alias」 フィールドに表示できます。

Activate Item Definitions

ソース・データからデータ・レンズで作成された品目定義をアクティブにするには、このチェック・ボックスを使用します。品目定義がアクティブな場合は、本番での属性抽出および標準化に品目定義を使用できます。品目定義がアクティブでない場合は、品目定義で認識される品目は、QA または本番では処理されません。

AutoBuild によって作成される品目定義は完全でない可能性があります。そのため、本番用にデータ・レンズをチェック・インする前に品目定義を Knowledge Studio で厳密に確認する必要があります。確認せずに本番用にデータ・レンズをチェック・インすると、照合アプリケーションの実行時に、予想を超える数のレコードが返される場合があります。これが該当するのは、本番と QA の Oracle DataLens Server のみです。開発サーバーでは非アクティブ・フラグが無視されます。

デフォルトでは、AutoBuild によって作成された品目定義は非アクティブになるので、このチェック・ボックスは選択されません。

「Create Match Rules」チェック・ボックス

デフォルト照合ルールを作成する場合は、このチェック・ボックスを選択します。各品目定義のデフォルトの属性順序が自動的に作成されます。デフォルトでは、ワークシートの列に指定した順序でスコアリング属性のみが含まれます。標準化タイプ「Match_Attributes」に基づいて、「Exact_Match」および「Close_Match」の照合タイプが作成されます。

このチェック・ボックスを選択すると、「Define Exact Match」および「Define Close Match」コントロールがアクティブになり、次のように使用できるようになります。

- **Define Exact Match コントロール**

品目定義ごとの Exact_Match ルールに対する必須属性の数を定義します。

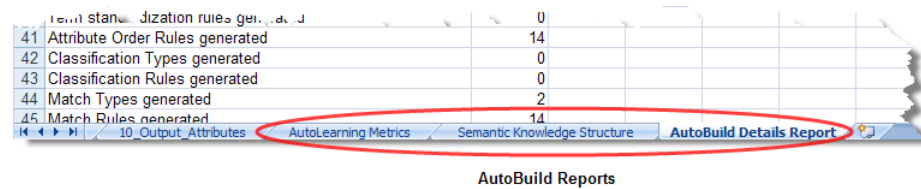
- **Define Close Match コントロール**

品目定義ごとの Close_Match ルールに対する必須属性の数を定義します。

AutoBuild レポート

AutoBuild アプリケーションでは、詳細なメトリック・データおよびデータ・レンズのセマンティック・ナレッジ構造を示す3つのレポートが自動的に生成されます。「Generate Report」オプションを選択すると、データ・レンズの作成後にこれらのレポートが作成されます。このオプションの設定の詳細は、2-23 ページの「Output Details」を参照してください。

レポートは Excel ワークブックに新規ワークシートとして追加されるので、情報を簡単に確認、印刷および保守できます。



AutoBuild を実行してデータ・レンズを生成するたびに、Excel ワークブックのレポートが更新されます。これは、ワークブック内のワークシートが1つでも複数でも行われます。

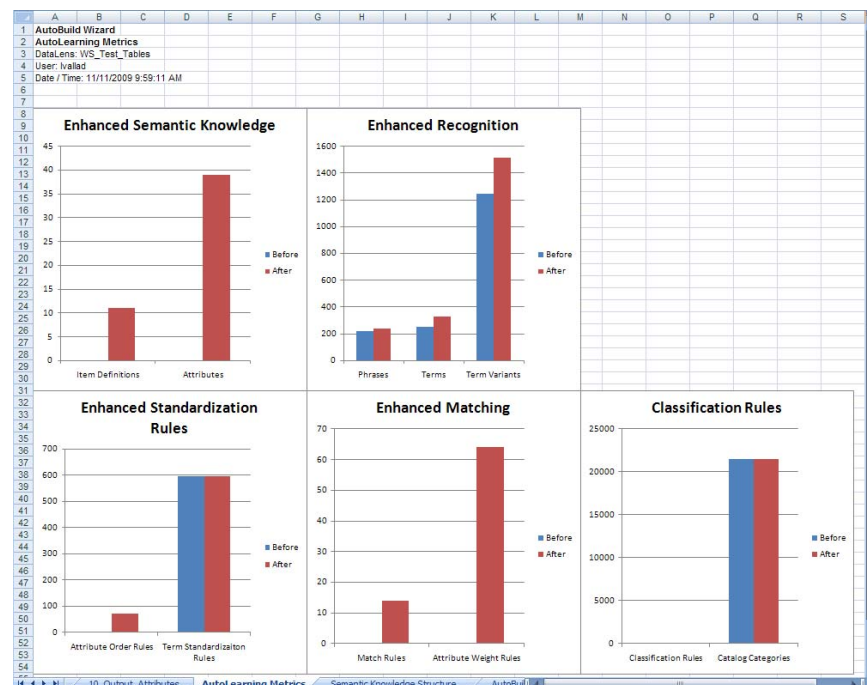
ヒント: 各ワークシートの名前を変更すれば、レポートが上書きされるのを回避し、将来の使用のために保持しておくことができます。

AutoBuild を使用して、これらのレポートを削除できます。「DataLens Tools」ボタンをクリックして、「Remove AutoBuild Reports」を選択します。

削除の確認を求めるプロンプトが表示され、「Yes」を選択すると、3つの AutoBuild レポートすべてがアクティブ・ワークブックから削除されます。

「AutoLearning Metrics」レポート

このレポートには、AutoBuild によってデータ・レンズに作成されたすべてのオブジェクトの概要が表示されます。次のサンプル・レポートに示すように、このレポートには、最も重要な各オブジェクトの前と後の状態を示す棒グラフが含まれています。



Semantic Knowledge Structure

次の例に示すように、このレポートには、AutoBuild で作成されたデータ・レンズを構成する品目定義、フレーズ、用語および代替カタログの詳細が表示されます。

A	B	C	D	E	F	G	H
1	AutoBuild Wizard						
2	Semantic Knowledge Structure						
3	DataLens: WS_Test_Tables						
4	User: lvalad						
5	Date / Time: 11/11/2009 9:59:12 AM						
6							
7							
8							
9	Item Definitions						
10	Item Definitions with Attributes						
11	Phrases						
12	Terms						
13	Alternate Catalog						
14							
15							
16							
17							
18							
19	Item Definitions						
20				New Meta-Data			
21				Renamed Meta-Data			
22				Existing Meta-Data			
23				Not Present in Meta-Data			
24							
25	Level: 1	Level: 2	Level: 3				
26	Writing instruments	Pens and pen refills					
27			Pen refills				
28			Pens				
29							
30		Markers					
31			China markers				
32			Highlighters				
33			Multi use markers				
34		Pencils					
35			Mechanical pencils				
36			Wooden pencils				
37							
38							
39							
40							
41	Item Definitions with Attributes						
42				New Meta-Data			
43				Renamed Meta-Data			
44				Existing Meta-Data			
45				Not Present in Meta-Data			
46							
47	Writing instruments						
48		Priority	Attribute				Phrase/Term Associations
49							
50	Pens and pen refills	Priority	Attribute				Phrase/Term Associations
51							
52	Pen refills	Priority	Attribute				Phrase/Term Associations
53							
54							

レポートの最初のリンクを使用すると、リンクに示されている特定のデータにナビゲートできます。

AutoBuild Details Report

このレポートには、データ・レンズの作成時に AutoBuild で実行されたすべてのアクションの詳細が表示されます。次のサンプル・レポートに示すように、これらのアクションには、データ・レンズの作成時刻、フレーズと用語と分類の改善、および AutoBuild によって収集された既存の警告が含まれます。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	AutoBuild Wizard													
2	AutoBuild Details Report													
3	DataLens: WS_Test_Tables													
4	User: Invalid													
5	Date / Time: 11/11/2009 9:59:13 AM													
6														
7														
8														
9	Totals													
10	Performance													
11	Phrase Refinements													
12	Term Refinements													
13	Classification Refinements													
14	AutoBuild Warnings													
15														
16														
17														
18														
19														
20														
21	Totals													
22														
23	Input Meta-Data Metrics													
24	Unique Categories found		11											
25	Leaf Categories found		7											
26	Attributes found		39											
27	Attribute values found		603											
28	Category Depth		3											
29														
30														
31	AutoLearning Metrics													
32	New Item Definitions		11											
33	New Attributes		39											
34	New Phrases		19											
35	New Phrase Productions		104											
36	New Terms		78											
37	New Term Variants		272											
38	Standardization Types generated		1											
39	Term variant standardization rules generated		0											
40	Term standardization rules generated		0											
41	Attribute Order Rules generated		14											
42	Classification Types generated		0											
43	Classification Rules generated		0											
44	Match Types generated		2											
45	Match Rules generated		14											
46	Category Sample Records generated		0											
47	Category Sample Files generated		0											
48	Random Sample Records generated		99											
49	Random Sample Files generated		1											
50			0											
51														
52														
53														
54														

レポートの最初のリンクを使用すれば、リンクに示されている特定のデータにナビゲートできます。「AutoBuild Warnings」セクションを確認し、データ・レンズを調整し改善してください。

AutoBuild のプロセス

この章では、品目情報を含む構造化データを AutoBuild で読み込んで、Knowledge Studio で使用できるデータ・レンズを生成する方法を最初から最後まで例で示します。ここでは、ペンや鉛筆のような標準の筆記具のいくつかのレコードを含むサンプル構造化データ・セットを使用します。

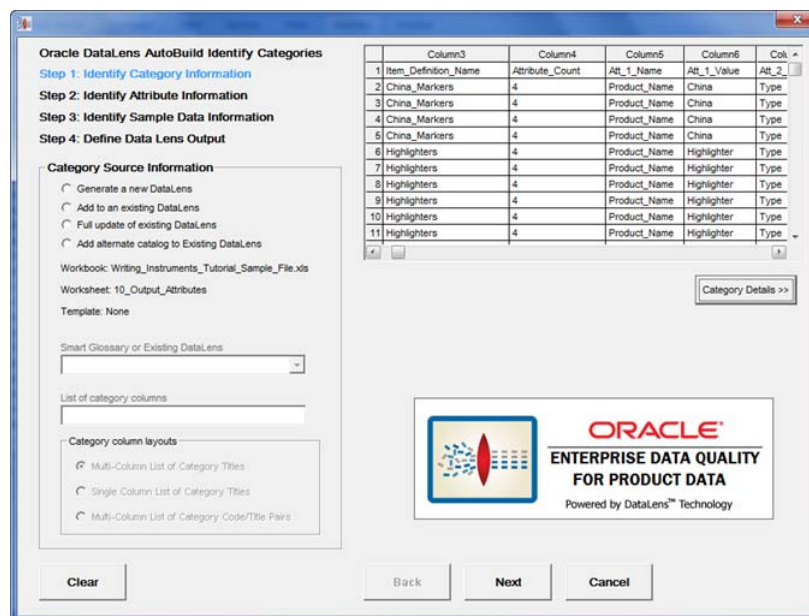
AutoBuild プロセスの使用

AutoBuild プロセスは、構造化データを含む Excel ワークブック・ファイルをオープンして開始します。ワークブックがオープンしたら、Services for Excel のアドイン・ツールバーを探します。インストールまたは AutoBuild ウィザードの詳細は、2-2 ページの「AutoBuild のユーザー・インタフェース」を参照してください。

id	Description	Level 1	Level 2	Level 3	Product_Name	Type	Ink_color
2							
3	W17 Ballpoint Pen Refill Fine Point Blue Ink 2 / Pack	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Balpoint	Blue Ink
4	W18 Ballpoint Pen Refill Med Pt Black Ink 2 / Pk	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Balpoint	Black Ink
5	W19 Ballpoint Pen Refill Med. Point black, ink 2 / Pack	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Balpoint	black ink
6	W110 Ballpoint Pen Refill Mdm Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Balpoint	Black Ink
7	W111 Ballpoint Pen Refill mdm point 7mm Ink Black	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Balpoint	ink Black
8	W119 Countertop pen refill Medium Point black ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Balpoint	black ink
9	W122 Security Pen Refills Medium Point Blue Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Balpoint	Blue Ink
10	W123 Counter Pen Refills Medium Point Blue Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Balpoint	Blue Ink
11	W137 Gel Grip Rollerball Pen Refill 11MM Red	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Rollerball	Red Ink
12	W138 Gel Ink Roller Ball Pen Refills 7MM Black Ink 2 / Pack	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Roller Ball	Black Ink
13	W142 Gel Ink Roller Pen Refill Fine Point Orange Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Roller	Orange Ink
14	W143 Roller Pen Refill Fine Point Violet Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Roller	Violet Ink
15	W181 Pen Refill Fine point Blue 2 / Pack	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill		Ink Black
16	W182 Pen Refill Ink Black 2	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill		
17	W183 Pen Refill Bunt Med Bk 5 / Bx	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Rfl	Bpnt	
18	W184 Pen Refill Med Bk 10/Bx	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Rfl		
19	W193 Refill Rollerball Ink Black	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Rollerball	Ink Black
20	W194 Refill Rollerball Ink Blue	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Rollball	Ink Blue
21	W1103 Roller Ball Pen Refill Medium Point Blue Ink	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	Roller Ball	Blue Ink
22							
id	Description	Level 1	Level 2	Level 3	Product_Name	Type	Ink_color
23							
24	W16 Ballpoint Pen Medium Point Black Barrel Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Balpoint	Black Ink
25	W117 Comfort Grip Retractable Ballpoint Pen medium Point Blue Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Retractable Ballpoint	Blue Ink
26	W118 Comfort Grip Retractable Ballpoint Pen Medium Point Red Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Retractable Ballpoint	Red Ink
27	W127 Retractable Ballpoint Pen Non-Refillable Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Retractable Ballpoint	Black Ink
28	W128 Erasable Ink Ballpoint Pen Medi Point Blue Barrel/Blue Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Balpoint	Blue Ink
29	W129 Felt Tip Pen Extra-Fine Point Translucent Blue Barrel Purple Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Felt Tip	Purple Ink
30	W130 Felt Tip Pen Extra-Fine Point Translucent black Barrel Red Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Felt Tip	Red Ink
31	W179 Pen Extra-Fine Point Blue Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen		Blue Ink
32	W180 Pen Extra-Fine Point Green Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen		Green Ink
33	W188 Porous Pen Extra-Fine Point Red Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Porous	Red Ink
34	W191 Porous Pen Extra-Fine Tip White Barrel/Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Porous	Black Ink
35	W195 Ballpoint Pen Retractable Fine Point Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Balpoint	Black Ink
36	W196 Retractable Ballpoint Pen Fine Point Clear Barrel/Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Retractable Ballpoint	Black Ink
37	W197 Retractable Ballpoint Pen Medium Point Black Barrel/black ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Retractable Ballpoint	black ink
38	W198 Retractable Ballpoint Pen Refillable Medium Point Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Retractable Ballpoint	Black Ink
39	W199 Ballpoint Pen Rubber Grip Fine Point Retractable Black Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Retractable Ballpoint	Black Ink
40	W1100 Retractable Ballpoint Pen Rubber Grip Fine Point Red Ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Retractable Ballpoint	Red Ink
41	W1101 Roller Ball Pen Gel Ink Refillable Medium Blue	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Roller Ball	
42	W1102 Roller Ball Pen Refillable Medium Red gel ink	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Roller Ball	
43	W1107 Pen , Balpoint , Blue Ink, Medium Point	Writing_Instruments	Pens_and_Pen_Refills	Pens	Pen	Balpoint	Blue Ink
44							
id	Description	Level 1	Level 2	Level 3	Product_Name	Type	Ink_color
45							
46	W112 China Mkrk Green Ink 12 / Box	Writing_Instruments	Markers	China_Markers	China	Mkrk	Green Ink
47	W113 China Markers red ink	Writing_Instruments	Markers	China_Markers	China	Markers	red ink
48	W114 China Markers White ink	Writing_Instruments	Markers	China_Markers	China	Markers	White ink
49	W115 China Mkrk yellow ink	Writing_Instruments	Markers	China_Markers	China	Mkrk	yellow ink
50							
id	Description	Level 1	Level 2	Level 3	Product_Name	Type	Ink_color
51							
52	W116 Chisel Point Highlighter Turquoise Ink	Writing_Instruments	Markers	Highlighters	Highlighter	Chisel Point	Turquoise Ink
53	W131 Highlighter Chisel Tip Fluorescent Purple	Writing_Instruments	Markers	Highlighters	Highlighter	Chisel Tip	Fluorescent Purple
54	W132 Highlighter Chisel Tip Fluorescent Yellow	Writing_Instruments	Markers	Highlighters	Highlighter	Chisel Tip	Fluorescent Yellow

Excel を使用して構造化データ・ファイルを作成する場合は、Excel では特定のデータが式や日付などのための Excel 固有の書式として解釈されることに注意することが重要です。すべてのデータ・セルの先頭に一重引用符 (') を使用して、データが厳密にテキストとして解釈されるようにしておくことをお勧めします。先頭に一重引用符があると、Excel では、セルの情報がテキストとして扱われます。たとえば、スラッシュ、ダッシュ、先頭の空白、ピリオドのある部品番号がセルに含まれている場合があります。

「AutoBuild」 ボタンをクリックするか、「DataLens Tools」メニューを選択して、「AutoBuild」 ボタンをクリックします。



AutoBuild アプリケーションの初回開始時のウィザードの表示では、デフォルトでアクティブな Excel ワークシートがデータソースとして選択されています。AutoBuild ウィザードでは、Excel をオープンしている間は、4つのウィザード・ステップで選択した設定が記憶されています。ワークブックを保存しておく、設定は永続的に保存されます。最初のステップに戻って「Clear」をクリックすれば、現在選択されているオプションをいつでもクリアできます。

カテゴリ情報の特定

構造化データのどの列にカテゴリ情報が含まれているかを特定します。これらのカテゴリ列は、AutoBuild がデータ・レンズの品目定義を作成するために使用され、データ・レンズのナレッジはカテゴリ階層に自動的に同期されます。品目定義階層は、ワークシート内のカテゴリ列のレイアウト、またはその他の分類参照ファイル（UNSPSC 業界標準など）に定義されている階層のどちらかで定義されます。

「List of Category Columns」の選択

選択したワークシートのカテゴリ列を特定するには、「List of category columns」フィールドに列のリストを入力します。カテゴリ列には、単一の列番号、「1-10」のような列範囲、または「1,5-10,16」のようにカンマで区切った列番号と範囲の組合せを指定できます。

カテゴリ列のリストを入力してから、[Tab] キーを押すか別のコントロールをクリックすると、「Category/Attribute Data」表に情報が再表示されます。選択した列の色が緑色になります。

	Column3	Column4	Column5	Column6	
1					
2	Level 1	Level 2	Level 3	Product_Name	T
3	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
4	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
5	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
6	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
7	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
8	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	
9	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	
10	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	
11	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	F

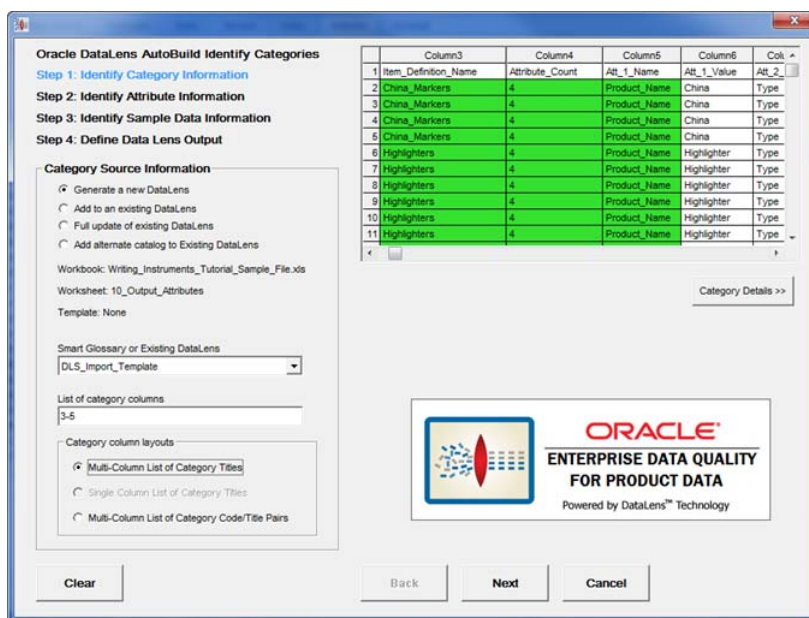
「Category Column Layout」の選択

3つのカテゴリ列レイアウトから選択できます。1番目の、最もよく使用されるオプションは、「**Multi-Column List of Category Titles**」オプションです。この列レイアウトでは、カテゴリ階層に複数のレベルがあり、各列が階層の1つのレベルを表していると見なされます。AutoBuildでは、カテゴリ階層が左から右へと表され、左の最初の列がカテゴリ階層の最高レベル、右の最後の列が階層の最低レベルを表すと見なされます。列リストに複数の列が含まれる場合は、常にこのレイアウトがデフォルトとして選択されます。

2番目のカテゴリ・レイアウトは「**Single Column List of Category Titles**」オプションです。1つのカテゴリ列のみをリストすると、このレイアウトが自動的に選択されます。

3番目のカテゴリ・レイアウトは「**Multi-Column List of Category Code/Title Pairs**」です。このレイアウトは先に説明した複数列リストに似ていますが、各カテゴリが列のペアとして表されます。最初の列にはカテゴリ・コード、2番目の列にはカテゴリ名が含まれています。

この例では、自動的に選択される「**Multi-Column List of Category Titles**」オプションを使用します。



「Next」をクリックして、次のステップに進みます。

属性情報の特定

構造化データで見つかった属性および属性値情報の場所とレイアウトを特定します。この属性情報は、AutoBuild でデータ・レンズの品目定義を作成するために使用され、そのデータ・レンズの名前はビジネス・ドメイン・スキーマに自動的に同期されます。属性値情報は、AutoBuild でテキスト自動認識用のデータ・レンズのフレーズおよび用語を作成するために使用されます。さらに、AutoBuild では、データ・レンズの適切なフレーズが正しい品目定義に自動的に関連付けられ、属性抽出が自動的に行われます。

必須属性情報の選択

まず、品目定義の必須属性として使用する属性列を指定します。通常、これらの必須属性列は、品目名、タイプおよびブランドで構成されています。

この例では、列 6 に製品名が次のように含まれています。

	Column3	Column4	Column5	Column6	
1					
2	Level 1	Level 2	Level 3	Product_Name	T
3	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
4	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
5	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
6	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
7	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
8	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
9	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
10	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B
11	Writing_Instruments	Pens_and_Pen_Refills	Pen_Refills	Refill	B

各必須ロジック列の指定を入力すると「Category/Attribute Data」表が更新されて、次に示すように、選択した列が紫色になります。

	Column4	Column5	Column6	Column7
1				
2	Level 2	Level 3	Product_Name	Type
3	Pens_and_Pen_Refills	Pen_Refills	Refill	Ballpoint
4	Pens_and_Pen_Refills	Pen_Refills	Refill	Ballpoint
5	Pens_and_Pen_Refills	Pen_Refills	Refill	Ballpoint
6	Pens_and_Pen_Refills	Pen_Refills	Refill	Ballpoint
7	Pens_and_Pen_Refills	Pen_Refills	Refill	Ballpoint
8	Pens_and_Pen_Refills	Pen_Refills	Refill	
9	Pens_and_Pen_Refills	Pen_Refills	Refill	
10	Pens_and_Pen_Refills	Pen_Refills	Refill	
11	Pens_and_Pen_Refills	Pen_Refills	Refill	Rollerball

構造化データ列は、複数の目的に使用される場合があります。たとえば、多くの場合、名前とタイプ情報は、カテゴリ情報と重なっています。

スコアリング属性情報の選択

次の例に示すように、カテゴリ情報のすぐ右側に、列のセットとして、属性情報がレイアウトされます。

Column6	Column7	Column8	Column9	Column10	Column11	Column12	
1							
2	Product_Name	Type	Ink_color	Size	Ink_Type	Barrel_Color	Grip_Ty
3	Refill	Ballpoint	Blue Ink				
4	Refill	Ballpoint	Black Ink				
5	Refill	Ballpoint	black, ink				
6	Refill	Ballpoint	Black Ink				
7	Refill	Ballpoint	ink Black	.7 mm			
8	Refill		black Ink				
9	Refill		Blue Ink				
10	Refill		Blue Ink				
11	Refill	Rollerball		1 MM			Gel Grip

この例では、列7から13が属性スコアリング情報として使用されます。

属性スコアリングを追加するには、「Add Scoring Attributes」チェック・ボックスを選択します。この例では、ヘッダー行に名前があるので、「Attribute names in the category file header」オプションを選択する必要があります。

次に、「Scoring Attribute」フィールドに属性列リストを入力する必要があります。カンマ区切りの列番号および列の範囲を指定できます。

属性列リストが受け入れられると、以前に説明したように、選択したスコアリング属性が青色になります。

Column5	Column6	Column7	Column8	Column9	Column10	
1						
2	Level 3	Product_Name	Type	Ink_color	Size	Ir
3	Pen_Refills	Refill	Ballpoint	Blue Ink		
4	Pen_Refills	Refill	Ballpoint	Black Ink		
5	Pen_Refills	Refill	Ballpoint	black, ink		
6	Pen_Refills	Refill	Ballpoint	Black Ink		
7	Pen_Refills	Refill	Ballpoint	ink Black	.7 mm	
8	Pen_Refills	Refill		black Ink		
9	Pen_Refills	Refill		Blue Ink		
10	Pen_Refills	Refill		Blue Ink		
11	Pen_Refills	Refill	Rollerball		1 MM	

次に選択例を示します。

「Next」をクリックして、次のステップに進みます。

サンプル・データ情報の特定

データ・レンズを作成する際に AutoBuild で使用できる、使用可能なサンプル・データの場所とレイアウトを特定する必要があります。このサンプル・データ情報は、AutoBuild で、生成するデータ・レンズにカテゴリ・ベースのサンプル・データ・ファイルと 100 レコードのランダム・サンプル・データ・ファイルの両方を追加するために使用されます。ランダム・サンプル・データ・ファイルでは、データ・レンズ・ナレッジの収束をデータ・アナリストが決定できます。カテゴリ・ベースのファイルでは、コンテンツ・エキスパートが、特定の品目商品に焦点を合わせて簡単に保守を実行できます。サンプル・データを使用できない場合は、使用可能な属性情報からサンプルを作成するオプションを使用できます。

サンプル・データ・ソースの選択

まず、サンプル・データ情報のデータ・ソースを選択します。Knowledge Studio を使用してサンプル・データを収集することも、ウィザードの前のステップで指定した属性情報からサンプル・データを作成することも、別のファイルからサンプル・データを選択することもできます。

この例では、属性からサンプルを作成しているので、適切なオプションおよび「Prefix Sample Data Columns (listed below)」チェック・ボックスが選択されています。

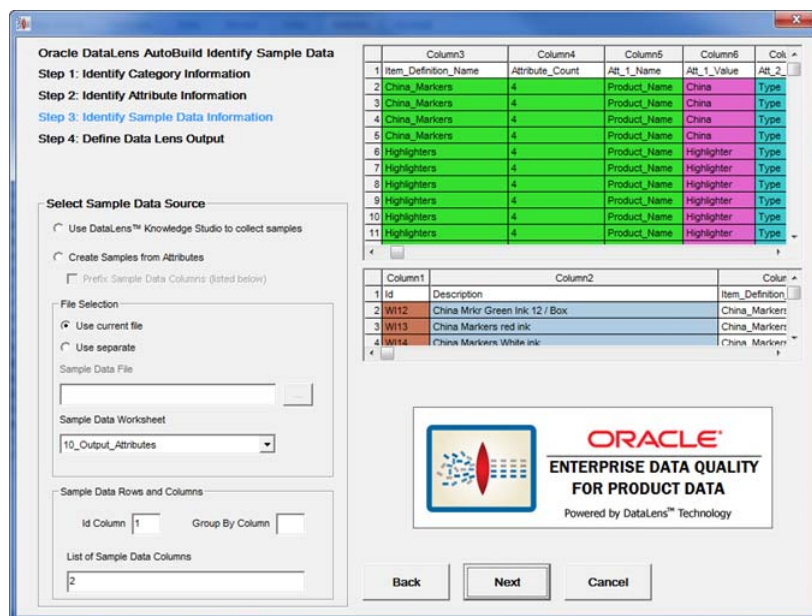
これらを選択すると、Knowledge Studio で使用するサンプル・データの説明が、使用可能な属性情報から自動的に作成されます。これは、サンプル・データの説明を手でできない場合に非常に便利です。さらに、生成される説明には、構造化データ・ソースの他の列が接頭辞として付加されます。この例では、名前、タイプおよびブランド名の列が、生成される説明に含まれています。

指定した属性情報から効果的にサンプルの説明を生成するには、一意のキーを使用して属性情報をリンクする必要があります。これらの列は列 5 から 7 で、この情報が「List of Sample Data Columns」フィールドに入力されています。

この例では、オープン中のワークブックの「10_Output_Attributes」ワークシートのソース・データが使用されています。

この例では、ID 番号が一意のキーとして使用されており、その列番号が「Id Column」フィールドに入力されています。

次に選択例を示します。

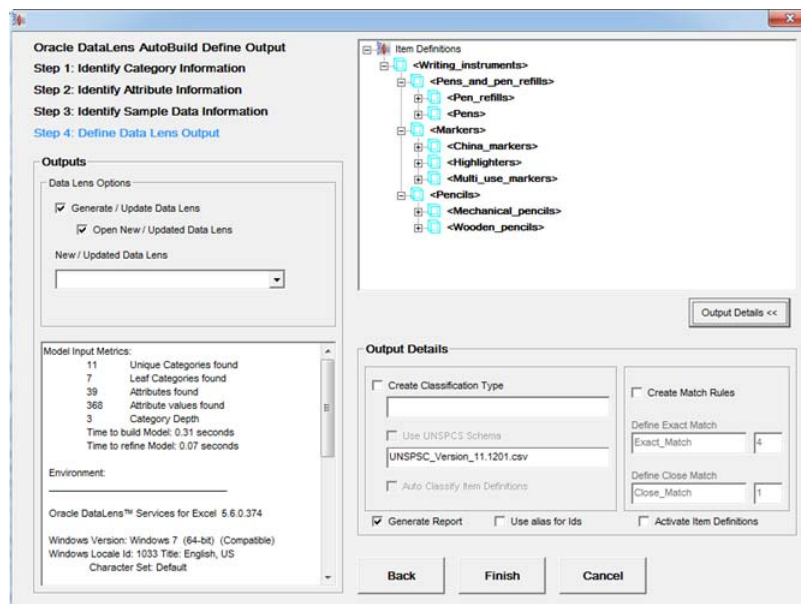


「Next」をクリックして、最後のステップに進みます。

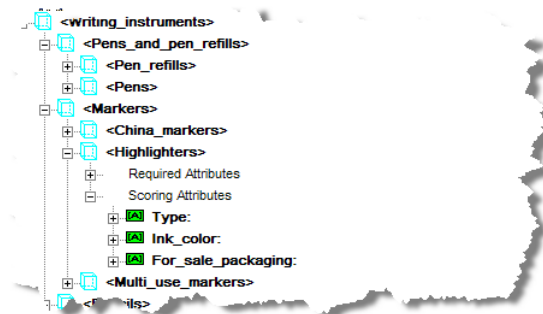
先に進む前に、AutoBuild では、ウィザードを使用して指定したカテゴリ情報と属性情報が読み取られ、その結果として得られる品目定義、品目定義属性、属性とフレーズの関連付け、および関連するフレーズおよび用語を含むモデルが生成されます。

データ・レンズ出力の定義

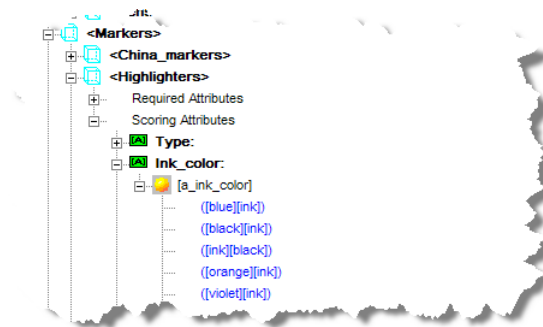
ウィザードの最終ステップでは、データ・レンズの生成前に、作成されるデータ・レンズの構造の確認、名前付け、および特定のオプションの選択を行います。



品目定義およびそれに関連する属性を確認して、期待どおりの結果がデータ・レンズに作成されることを確認します。この例では、「Writing Instruments」品目定義に、子品目定義全体が表示されます。「Highlighters」を展開すると、これらの属性を「Scoring Attributes」で次のように表示できます。

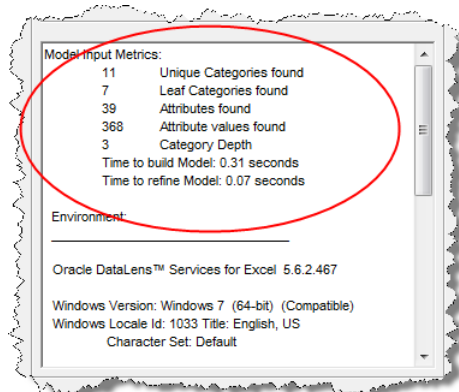


さらにドリル操作を行えます。「Ink_color」属性を展開すると、「[a_ink_color]」フレーズで、フレーズ生成結果全体を次のように表示できます。



品目定義階層が期待どおりでないことがわかった場合は、「Back」ボタンを選択して AutoBuild の前のダイアログに戻り、選択をやり直します。

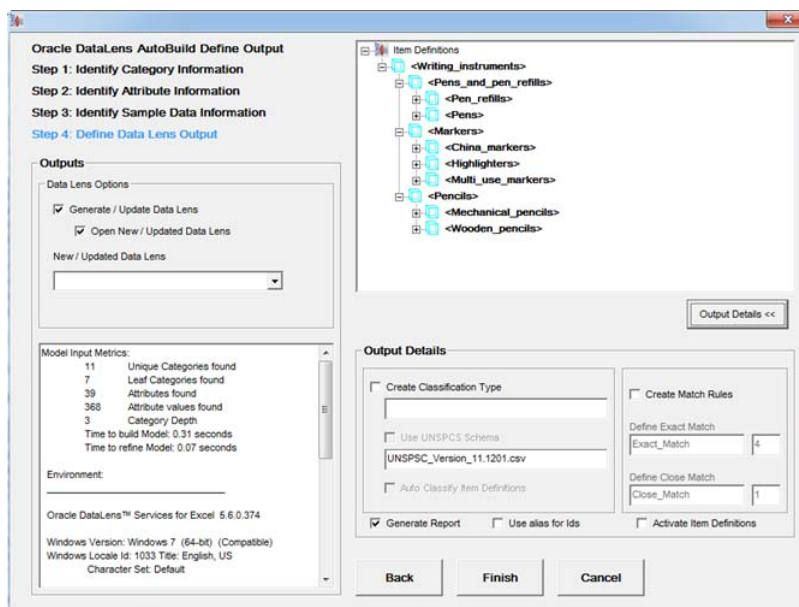
ステータス・ペインのメッセージを確認して、レポートされているメトリックが正しく、警告メッセージやエラー・メッセージがないことを確認します。



通常 AutoBuild では、警告やエラーが表示されても、新規データ・レンズの作成を続行できます。多くの場合、構造化データの例に重複または競合するカテゴリ情報が見つかったことが報告されます。警告を確認して、構造化入力を訂正する必要があるかどうかを決定してください。

注意：「DLS_」という接頭辞は、データ・レンズがスマート用語集であることを示すために予約されているので、データ・レンズ名には使用しないでください。

ウィザードの最終ステップの選択例を次に示します。

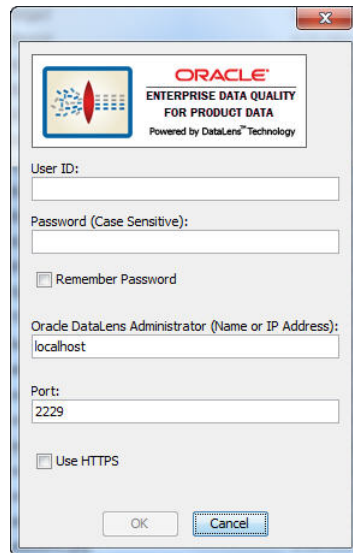


AutoBuild プロセスを完了して、新規レンズを作成するには、「Finish」をクリックします。構造化データに指定されたカテゴリ情報と属性情報から新規データ・レンズが作成されます。この例では、「Open New / Updated Data Lens」オプションが設定されていたので、AutoBuild プロセスが完了すると、Oracle Enterprise Data Quality for Product Data のログイン・ダイアログ・ボックスが表示されます。



ユーザー名とパスワードを入力します。「Remember Password」チェック・ボックスを選択すれば、ログオンするたびにパスワードを入力する必要がなくなります。

Oracle DataLens Server を変更する場合、または HTTP Secure (HTTPS) を使用する場合は、「Change Server」をクリックします。HTTPS オプションの動作が保証されるのは、WebLogic をアプリケーション・サーバーとして使用する Oracle DataLens Server のみです。



Oracle DataLens Server を変更するには、Oracle DataLens Administration Server のホスト名または IP アドレスおよびポート番号を入力します。

注意： Oracle DataLens Server が WebLogic アプリケーション・サーバーを使用している場合は、このポート番号が WebLogic ドメイン・サーバーのポート番号と一致している必要があります。

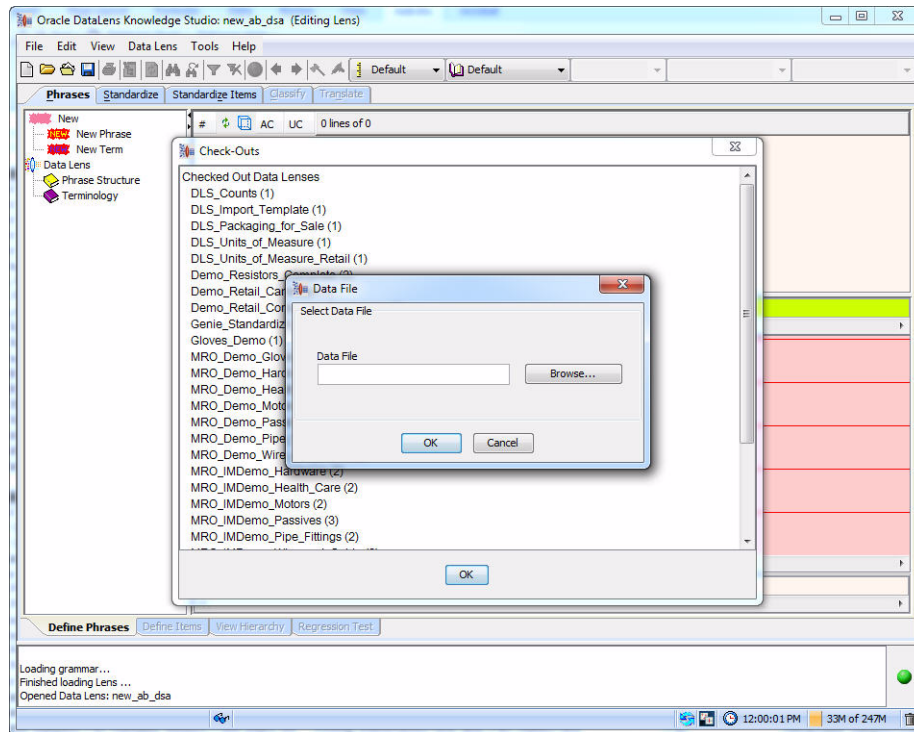
Oracle DataLens Server との通信に HTTPS を使用するには、「Use HTTPS」チェック・ボックスを選択します。

注意： Oracle DataLens Server (および EDQP トポロジ内の他のすべてのサーバー) を、すべて HTTPS を使用するように構成しておく必要があります。HTTPS の構成および使用の詳細は、『Oracle Enterprise Data Quality for Product Data Oracle DataLens Server 管理者ガイド』を参照してください。

すべての情報が正しければ、「OK」をクリックします。

Oracle Enterprise Data Quality for Product Data Launch Pad が表示され、Knowledge Studio が自動的に開始されて、サンプル・ファイルの選択を求めるプロンプトが表示されます。

注意： Oracle Enterprise Data Quality for Product Data Launch Pad の「Oracle DataLens Knowledge Studio」ボタンをクリックすると、Knowledge Studio が応答しなくなりオープンしなくなるので、クリックしないでください。



使用するサンプル・ファイルを選択して、「OK」をクリックすると、データ・レンズがオープンして AutoBuild の結果を表示できます。

これで、アプリケーション要件を満たすようにデータ・レンズを変更および拡張する準備が整いました。

新規データ・レンズの完成

AutoBuild によって生成されたデータ・レンズには、データ・レンズの開発を完了するために必要な構造が含まれています。AutoBuild によって生成された品目定義、属性、フレーズおよび用語は、開発プロセスの最終ステップのガイドとしての役割を果たします。先へ進む前に、生成されたレンズ構造を十分に理解してください。

データ・レンズを完成させる際には、生成されたコンポーネントをできるだけ再使用して、構造化品目情報の当初の意図との整合性を保つようにします。

Knowledge Studio を使用して新規データ・レンズを完成するために実行するステップについて、概要を次に示します。

1. 品目定義認識ロジックを完成します。

識別可能な品目名または品目タイプの情報が使用できない場合に、これは特に重要です。品目定義認識ロジックを完成すると、初期認識のあいまいさの多くが Knowledge Studio によって解決されます。これによって、残りのステップも正しく行えるようになります。

2. データ・レンズのフレーズ・レベルのあいまいさをすべて調べます。

各標準ランダム・サンプル・ファイルを確認して、このような複数の参照を削除するか解決します。

3. 標準サンプル収束を使用して、品目認識を完成します。

標準のデータ・レンズ開発方法を適用して、ランダム・サンプル・ファイルを使用したデータ・レンズ認識を完成します。

4. 標準化ルールを追加します。

データ品質基準および開発中のデータ・アプリケーションのどちらかまたは両方に必要な標準化ルールを作成します。AutoBuild プロセスの標準スマート用語集を使用すれば、この作業を最小限に抑えることができます。スマート用語集を使用すれば、グローバルなすべての用語およびフレーズを即時に標準化できます。スマート用語集を使用して、すべての単位変換を前もって行うこともできます。このレンズに固有の用語については、「Standardize」タブの「Copy Full Form」機能を使用すれば、標準化作業を効率的に開始できます。

用語とフレーズの標準化を完了したら、出力に必要な属性を選択し、必須照合ルールを設定して、標準化作業を終了できます。

注意：「Create Replace all Rules」オプションを選択した場合は、その前の標準化が自動的に行われます。

5. レンズの回帰テスト・セットを作成および検証します。

Knowledge Studio を使用して、レンズの回帰テスト・セットを作成し、将来の保守作業を正しくテストできるようにします。初回の回帰セットを注意深く確認して、結果がすべて期待どおりであることを確認し、必要に応じて調整する必要もあります。

6. 新規レンズをチェック・インし、開発用にデプロイし、アプリケーションのテスト・ケースに対してテストします。

Application Studio で作成された各データ・アプリケーションは、レンズが追加または更新されたときに常に実行する必要がある関連テスト・ケースを持ちます。

7. AutoBuild の構成を改善します。

これまでのステップの結果を使用して、新規データ・レンズの作成に使用する AutoBuild のスマート用語集を更新および改善します。可能な場合は、処理結果を使用して構造化品目情報のコンテンツを改善します。AutoBuild で使用される情報を改善すると、AutoBuild で生成されるデータ・レンズも明らかに改善します。

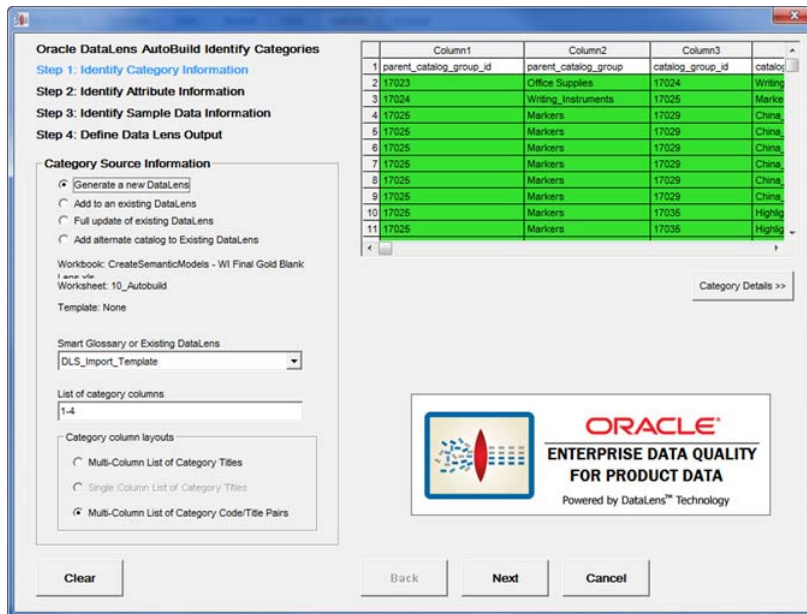
データ・ペアのデータ・レンズの例

この項では、属性名と変数の両方が各レコードに含まれているデータを使用する例を説明します。このペアになったデータは、「AutoBuild プロセスの使用」の項で説明した方法と同じ方法で AutoBuild プロセスで使用されていますが、「Attribute Information」のオプションが異なります。

ペアになったデータの Excel ワークシートの例を次に示します。

A	B	C	D	E	F	G
parent_catalog_group_id	parent_catalog_group	catalog_group_id	catalog_group	attr_id	attr_name	attr_value
17023	Office Supplies	17024	Writing Instruments			
17024	Writing Instruments	17025	Markers			
17025	Markers	17029	China Markers	1152 2530	Type	
17025	Markers	17029	China Markers	1152 2531	Ink_Color	Green
17025	Markers	17029	China Markers	1152 2531	Ink_Color	Red
17025	Markers	17029	China Markers	1152 2531	Ink_Color	Yellow
17025	Markers	17029	China Markers	1152 2531	Ink_Color	White
17025	Markers	17029	China Markers	1152 2532	For_Sale_Packaging	12
17025	Markers	17035	Highlighter	1153 2534	Type	Chisel Tip
17025	Markers	17035	Highlighter	1153 2534	Type	Chisel Point
17025	Markers	17035	Highlighter	1153 2535	Ink_color	Pink

カテゴリ情報を選択すると、次のように表示されます。



ソース・データがペアになっていることは、AutoBuild ウィザードのステップ 2 で次のように示します。

Oracle DataLens AutoBuild Identify Attributes

Step 1: Identify Category Information
 Step 2: Identify Attribute Information
 Step 3: Identify Sample Data Information
 Step 4: Define Data Lens Output

Attribute Information

Required Attributes

Add Required Attributes Single term columns

Item Name Column(s) Any Of 4

Item Type Column(s) All Of

Unique Name Column(s) Any Of

Use names and types to create required attributes

Scoring Attributes

Add Scoring Attributes

Attribute names are in the same row as the category

Row contains Attribute Names only

Row contains Attribute Name / Value pairs

Row contains Attribute Name/Value/UOM

Attribute names in the category file header

Scoring Attributes Single term columns

6,7

Optional Attributes

Add Optional Attributes

All Optional attribute columns (single term)

Column3	Column4	Column5	Column6	Column7
1 catalog_group_id	catalog_group	attr_id	attr_name	attr_value
2 17024	Writing_Instruments			
3 17025	Markers			
4 17029	China_Markers	1152 2531	Type	
5 17029	China_Markers	1152 2531	Ink_Color	Green
6 17029	China_Markers	1152 2531	Ink_Color	Red
7 17029	China_Markers	1152 2531	Ink_Color	Yellow
8 17029	China_Markers	1152 2531	Ink_Color	White
9 17029	China_Markers	1152 2532	For_Sale_Packaging	12
10 17035	Highlighter	1153 2534	Type	Chisel Tip
11 17035	Highlighter	1153 2534	Type	Chisel Poi

Attribute Details >>

ORACLE
 ENTERPRISE DATA QUALITY
 FOR PRODUCT DATA
 Powered by DataLens™ Technology

Back Next Cancel

この例では、スコアリング属性は列6と7で、属性名と値データのペアになっています。

Oracle DataLens AutoBuild Identify Sample Data

Step 1: Identify Category Information
 Step 2: Identify Attribute Information
 Step 3: Identify Sample Data Information
 Step 4: Define Data Lens Output

Select Sample Data Source

Use DataLens™ Knowledge Studio to collect samples

Create Samples from Attributes

Prefix Sample Data Columns (listed below)

File Selection

Use current file

Use separate

Sample Data File

Sample Data Worksheet

20_SampleData

Sample Data Rows and Columns

Id Column 2 Group By Column

List of Sample Data Columns

3

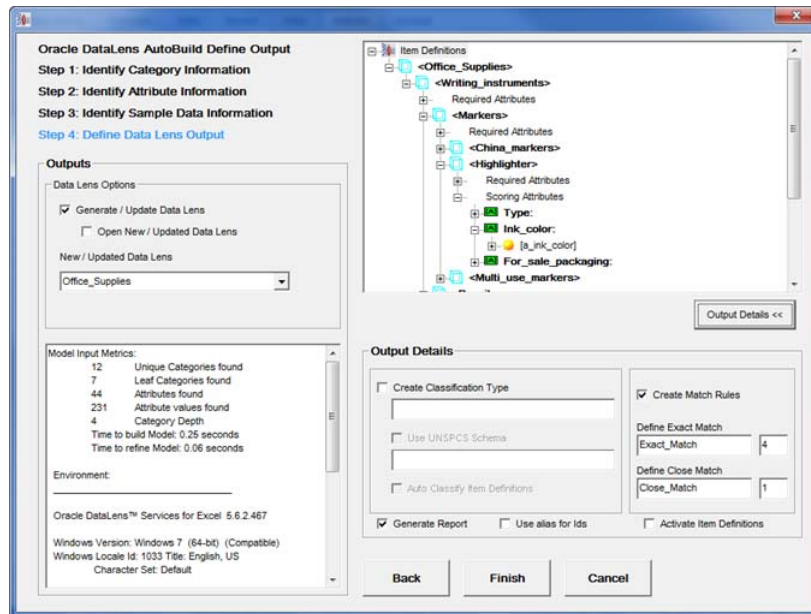
Column1	Column2	Column3	Column4
1 parent_catalog_group_id	parent_catalog_group	catalog_group_id	catalog_group_name
2 17023	Office Supplies	17024	Writing
3 17024	Writing_Instruments	17025	Markers
4 17025	Markers	17029	China
5 17025	Markers	17029	China
6 17025	Markers	17029	China
7 17025	Markers	17029	China
8 17025	Markers	17029	China
9 17025	Markers	17029	China
10 17025	Markers	17035	Highlighter
11 17025	Markers	17035	Highlighter

Column2	Column3	Column4
1 item_number	description	
2 1020	Highlighter Grip XL Chisel Tip Pink	
3 1014	Highlighter Chisel Point Blue	
4 1017	Highlighter Chisel Tin Blue	

ORACLE
 ENTERPRISE DATA QUALITY
 FOR PRODUCT DATA
 Powered by DataLens™ Technology

Back Next Cancel

ウィザードの最終ステップでは、「Step 4: Define Data Lens Output」セクションでペア・データがなかった場合と同様の構造で、品目定義が表示されます。



Excel ワークシートでは、様々な方法でデータを列に編成できます。AutoBuild のオプションの選択によって、結果のデータ・レンズに作成されるナレッジ構造が定義されます。

その他の考慮事項

AutoBuild アプリケーションを最大限に活用するために、次の点を確認することをお勧めします。

カテゴリ

カテゴリは、サンプルの説明、品目名および品目タイプに対して使用できるのに対し、カテゴリに対して使用できるのは品目名とタイプです。

既存のデータ・レンズの拡張

構造化品目情報から新規データ・レンズを作成して、必要に応じてデータ・レンズを調整します。「[Step 1: Identify Category Information](#)」の既存のデータ・レンズの追加または更新オプションのいずれかを使用して、AutoBuild を再実行します。

例の作成

適切に構造化された最適データがない場合、カテゴリごとに数個のサンプル・レコードを作成すれば、データ・レンズの開発時間を大幅に削減できます。さらに、データ・レンズ用語インポート機能を使用して、データ認識のギャップを埋めてください。

過大なカテゴリ名

一般に、カテゴリ名では品目名とタイプのみを使用します。追加情報が含まれているカテゴリ名は過大と見なされます。たとえば、「Surgical Gloves Latex-Free Powder-Free Nitrile Size 7」という過大なカテゴリ名は、「Surgical Gloves」にすれば過大でなくなります。

カテゴリ名を標準化するか、正しいカテゴリ・レベルを使用すれば、作業を削減でき、データ・レンズ属性を中心とした照合を最適化できます。

過大な属性名

カテゴリ名と同様に、属性名も過大にしないようにします。たとえば、「Heart Rate Monitor Operating Voltage」は「Voltage」にします。

AutoBuild を使用して、データ・レンズを作成する前に属性名を標準化すると、作業を削減でき、データ・レンズ属性を中心とした照合を最適化できます。

クライアント・ソフトウェアのインストール

Enterprise DQ for Product では、Java Web Start を使用して、クライアント・デスクトップのソフトウェアの現行バージョンを初回インストールおよび保守します。このプロセスでは、Oracle DataLens Server にアクセスし、接続を起動して、ソフトウェアをダウンロードする必要があります。

Enterprise DQ for Product クライアント・アプリケーションのダウンロードおよびインストールは、Oracle DataLens Server のインストール・ページを表示し、Java Web Start を使用して次のように行います。

1. Java SE Runtime Environment (JRE) 6 アップデート 21 がインストールされていることを確認します。次のサイトで、JRE をダウンロードして、インストール手順を入手できます。

<http://www.oracle.com/technetwork/java/javase/downloads/index.html>

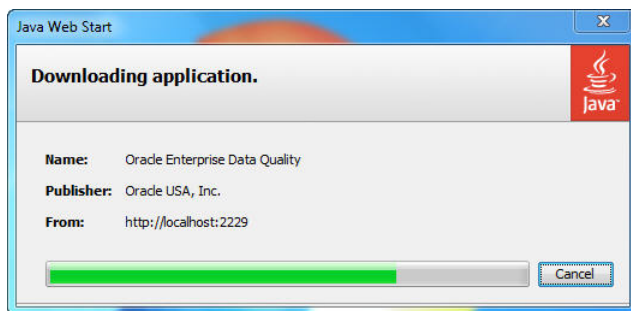
2. Microsoft Internet Explorer を起動します。
3. 次のサイトで、接続を起動してクライアント・ソフトウェアをダウンロードします。

`http://server:2229/datalens/datalens.jnlp`

ここで、*server* は、Oracle DataLens Server のホスト名を示しています。

注意： アプリケーション・サーバーのポート番号を 2229 以外に設定した場合は、Oracle DataLens Server に接続して、クライアント・アプリケーションをダウンロードする際に、上の URL でそのポート番号を使用する必要があります。

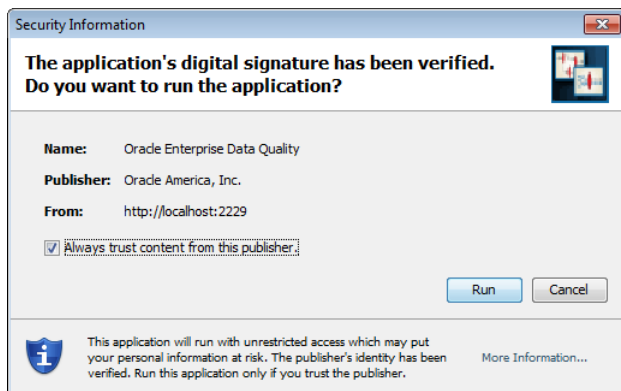
アプリケーションのダウンロードと検証が始まります。



注意： .jnlp ファイルがプログラムに関連付けられていないことを示すファイル・ダウンロード・メッセージを受信した場合は、サポートされている JRE がインストールされていません。「Cancel」をクリックしてステップ 1 に戻ります。

検証が完了した後、インストールが始まります。Enterprise DQ for Product のファイルは、信頼できる発行元によって電子署名されているので、次のセキュリティ警告が表示されま

す。



ヒント： 今後このセキュリティ・ダイアログが表示されないようにするには、「**Always trust content from this publisher**」チェック・ボックスを選択します。

4. 「Run」をクリックして続行し、インストールを完了します。

Oracle Enterprise Data Quality for Product Data のログイン・ダイアログが表示されます。

