

**Oracle® Enterprise Data Quality for Product Data**

用語集

リリース 5.6.2

部品番号 : B66255-01

2011 年 7 月

Oracle Enterprise Data Quality for Product Data 用語集, リリース 5.6.2

部品番号 : B66255-01

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.

---

---

# はじめに

この用語集は、「[関連ドキュメント](#)」に記載されている Oracle Enterprise Data Quality for Product Data（以前の Oracle Product Data Quality）のマニュアルで使用されている用語を定義することを目的としたものです。

## ドキュメントのアクセシビリティについて

オラクル社のアクセシビリティへの取り組みについての詳細情報は、Oracle Accessibility Program 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

詳細は、ドキュメント・セット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・ガイドには、Oracle Enterprise Data Quality for Product Data Oracle DataLens Server の詳細なインストール方法が記載されています。
- Oracle Enterprise Data Quality for Product Data Oracle DataLens Server 管理者ガイドには、Oracle DataLens Server のインストールと管理に関する情報が記載されています。
- Oracle Enterprise Data Quality for Product Data COM インタフェース・ガイドには、Oracle DataLens Server COM API のインストール方法と使用方法が記載されています。
- Oracle Enterprise Data Quality for Product Data Java インタフェース・ガイドには、Oracle DataLens Server Java API のインストール方法と使用方法が記載されています。
- Oracle Enterprise Data Quality for Product Data Application Studio リファレンス・ガイドには、Data Service Applications (DSA) の作成と保守について記載されています。
- Oracle Enterprise Data Quality for Product Data AutoBuild リファレンス・ガイドには、既存の製品情報とデータ・レンズのナレッジに基づいて最初のデータ・レンズを作成する方法が記載されています。
- Oracle Enterprise Data Quality for Product Data Knowledge Studio リファレンス・ガイドには、データ・レンズの実行と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 Services for Excel リファレンス・ガイドには、Microsoft Excel スプレッドシートに含まれるデータに基づいて DSA を作成する方法が記載されています。
- 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 アプリケーションで作成される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](http://download.oracle.com/docs/cd/E20593_01/index.htm)

## 表記規則

このマニュアルでは次の表記規則を使用します。

規則	意味
太字	太字は、操作に関連する Graphical User Interface 要素、または本文中で定義されている用語および用語集に記載されている用語を示します。
イタリック体	イタリックは、ドキュメントのタイトル、強調またはユーザーが特定の値を指定するプレースホルダ変数を示します。
固定幅フォント	固定幅フォントは、段落内のコマンド、URL、サンプル内のコード、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、ディレクトリ、パスの名前を示します。
固定幅フォント	太字の固定幅フォントは、入力するコマンドまたはテキストを示します。



---

---

# 用語集

## Application Studio

DSA を作成するための対話形式の設計システム。

## ASCII

英語で使用されている文字のみを対象とする標準の文字セット。

## AutoBuild アプリケーション

AutoBuild は、既存の Excel スプレッドシートの製品情報と Oracle Enterprise Data Quality for Product Data スマート用語集を使用して、エンタープライズ・コンテンツに固有の初期データ・レンズを作成するために使用するアプリケーションである。

## AutoLearn 機能

AutoLearn は、用語のバリエーションをそれに対応する、データ・レンズを構成する品目定義に關与する完全形式の用語に自動的に割り当てる、Oracle Enterprise Data Quality for Product Data の機能である。

## Data Service Application (DSA)

すべての内部的な統合および操作コンポーネントを呼び出しシステムから隠蔽する、Service Oriented Architecture (SOA) 技術に基づくソフトウェア・アプリケーション。DSA は、データを使用および操作してビジネスの問題を解決し、HTTP メッセージを使用して呼び出される。

DSA は、データ・レンズ間およびデータ・レンズ内を移動するときのコンテンツのフローと制御を定義する。ビジネス・プロセスを定義する。Data Service Application は、プロセス・マップとそれに関連する統合によって構成される。

## DataLens 手法

DataLens 手法は、エンタープライズ・データのドメインに関するコンテキスト・ナレッジの作成に基づいている。このコンテキストは、部品記述において見られるフレーズ構造ルールの定義によって取得される。

## eClass または eCl@ss

新しい標準化されたマテリアルおよびサービス分類。ドイツの先進企業によって開発された分類システム。サプライヤとカスタマの間の情報交換に関する標準として提供される。eCl@ss は、12,000 語のキーワードが登録された 4 レベルの階層分類システムという特徴を持つ。eCl@ss は、生産財購買者に対する市場構造を反映し、エンジニアによる開発、計画、保守をサポートする。階層またはキーワードを使用したアクセスにより、エキスパートと一時的ユーザーの両方が分類内を移動できる。詳細は、eCl@ss の Web サイトを参照してください。

<http://www.eclass.de>

## Governance Studio

ユーザーが容易に DSA プロジェクトを実行し、複数のグラフ化オプションで出力データを管理できるようにするクライアント・アプリケーション。

### ISO-8859-1 (Latin-1)

西欧言語（フランス語、ポルトガル語、イタリア語、ドイツ語、スペイン語、英語など）で使用される文字を対象とする ISO の標準キャラクタ・セット。Latin-1 キャラクタ・セットとも呼ばれる。詳細は、<http://www.htmlhelp.com/reference/charset> を参照してください。

### Knowledge Studio

エンタープライズ・データの標準化、分類、翻訳のためのツールとプロセスを提供するナレッジ・エンジニアリング製品。

### ODBC (Open Database Connectivity)

ネットワーク上のデータベースにアクセスするための共通言語をアプリケーションに提供するデータベース・プログラミング・インタフェース。

### Oracle DataLens Server

Oracle DataLens Server は、Oracle Enterprise Data Quality for Product Data Knowledge Studio やデータの無人バッチおよびトランザクション・プロセスに対して、自動化された標準化と分類のサービスを提供する。Oracle DataLens Server は、多数の対話型要求と同時実行ジョブの両方を処理できる。詳細なインストール方法およびセットアップ、構成、保守に関する情報は『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。

### SME

業務の専門家。企業データの特定の実体を詳細に理解している人。たとえば、会社の電子部品データベースを使用する電気エンジニアや製造エンジニア。

### Transformation Map

ユーザーが複数のナレッジ・ベースと他の情報の組合せの実行の順序を設定して複雑なセマンティック変換を実行できるようにするプロセス。

### UNSPSC

Universal Standard Products and Services Classification (UNSPSC)。商品を分類および識別するスキーマである。販売側と購入側のカタログで使用される。Electronic Commerce Code Management Association (ECCMA) は、UNSPSC コードの管理および開発を監督する非営利組織である。詳細は、ECCMA の Web サイトを参照してください。

<http://www.eccma.org>

### UTF-8

Universal Transformation Format 8 (キャラクタ・セット)。Oracle Enterprise Data Quality for Product Data によって使用されるテキスト文字のエンコード。このエンコードにより、Oracle Enterprise Data Quality for Product Data は世界中の国際キャラクタ・セットを扱うことができる。UTF-8 は Unicode の文字エンコード方式である。詳細は、Unicode Consortium の Web サイトを参照してください。

<http://www.unicode.org>

### XML

eXtensible Markup Language。XML は、構造化されたデータ・ファイルを作成するための構文である。HTML と同様に、XML ファイルにはファイルの編成に使用される一連のタグが含まれる。基本的な XML の構造は、1 対のタグとその間の内容で構成される。タグには、タグ付けの構造を変更する属性もある。詳細は、W3C の Web サイトを参照してください。

<http://www.w3.org/XML/1999/XML-in-10-points>

## エンタープライズ・データ

通常はバック・オフィス・システムからの、個別のレコードの形式を取るデータ。このデータは、通常は属性に富み、電子メールやメモ・ドキュメントで一般に見られるような自由形式のテキストではない。

「[明細項目](#)」を参照してください。

## オントロジ

一般的な意味としては、オントロジとは存在するものについての研究または考察。形而上学の1分野であり、物事の原則や本質の研究である。

情報技術においては、電子商取引や計画の「アクティビティ」など、知識や実践の特定のドメインにおけるエンティティと相互作用の作業モデルである。人工知能（AI）においては、「プログラムと人の共有知識を支援するために使用される概念化の仕様」である。この用法では、オントロジとは情報交換のための合意された語彙を作成するためになんらかの方法（特定の自然言語など）で指定される概念（物事、イベント、関係など）のセットである。

## 機械翻訳

自動化されたソフトウェア主導の技術を使用して、人が読むことのできるテキストを変換元の言語から変換先の言語に翻訳すること。

## クレンジング

「[標準化](#)」を参照してください。

## 構文

構文とは、言語ステートメント内の要素の文法、構造、または順序である。構文はコンピュータ言語にも自然言語にも適用される。通常、構文は「単語の順序」と考えられる。コンピュータ言語での構文は、ほとんどのアセンブラ言語でのように非常に厳格な場合や、任意の順序で記述できる「keyword」パラメータを使用する言語のようにあまり厳格ではない場合がある。

## コンテンツ

特定の種類の、または特定の対象についての、エンタープライズ・データの大規模なコレクション。そのようなコンテンツには重複するフレーズが大量に含まれることが多い。

## サンプル

企業データ・セットと同じように大多数の用語を表しているデータのランダム化されたコレクション。

## 情報サプライ・チェーン

データおよび情報の、物理的なサプライ・チェーンに相当するもの。この場合、システム間のデータ・フローにデータの変換および翻訳が伴う、システム間でのデータ（通常は製品データ）の移動である。通常、システム間の受け渡しは、手渡しまたはアドホックなツールで行われる。

## スマート用語集

スマート用語集とは、多くの種類の情報に共通する、または特定のドメインや業種の情報に固有である、フレーズと用語を認識するために構造化されているデータ・レンズである。品目の原材料や完成品に対して最も頻繁に使用されるフレーズおよび用語を含むスマート用語集は、一般的または同業種向けの水平レンズである。より具体的なアプリケーション（たとえば配管マテリアル）からのフレーズと用語を含むスマート用語集は、アプリケーション固有レンズまたは業種別の垂直スマート用語集である。

スマート用語集は、初期データ・レンズ・セットを迅速に作成するために AutoBuild アプリケーションによって最も頻繁に使用される。

## 正規表現

様々なテキスト形式を簡単な表現で取得する方法である。たとえば、すべての整数は `/\d+/` という正規表現パターンで表すことができる。正規表現構文の詳細は、[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。

## セマンティクス

意味を扱う記号の哲学または研究である。自然言語およびコンピュータ言語に関しては、構文（たとえば、文章での単語の順序や、正確なコンピュータ・コマンドの表記法）とセマンティクス（単語の本当の意味や、コマンドで要求されている機能）が区別されることがある。

## セマンティック変換

セマンティック・モデルを適用した結果である形式へのバック・オフィス・データのレコード・レベルでの変換。セマンティック変換の例としては、レコード記述の標準化、品目の分類、属性の抽出、言語の翻訳などがある。

## ソース書式設定

ルールを作成する前にコンテンツの書式を設定すること。書式設定の目的は、コンテンツの標準化に対する特殊ルールの数を減らすことである。

## 属性

目的の単一レコードに含まれる抽出可能な最小の情報単位。また、品目定義で定義されている、品目定義自体の完全性に関連している数量。品目の本質的なプロパティを説明する、品目の特徴である。

## データ・レンズ

データ・セット内の構造、コンテキスト、用語に関する情報を含むリポジトリ。アーキテクチャの観点からは、データ・レンズは、コンテキストの理解に関する情報が非常に少ない標準的なリレーショナル表とは大きく異なる。データ・レンズは、データから抽出されたコンテキスト・ナレッジを表すために Oracle DataLens Server によって使用される。

## ドメイン

特定の対象に関係していることが多い、文法ルールの任意のコレクション。2つのドメインには、重なり合う文法ルールまたは一緒に使用される文法規則が含まれる場合がある。

## トランザクション・プロセス

コンピュータ・プログラミングでは、通常、トランザクション・プロセスとは、要求を満たし、データとそれに関連するシステムの整合性を保証するために、1つの単位として扱われる、一連の情報交換および関連する作業を意味する。

## バック・オフィス・データ

バック・オフィスはほとんどの会社に存在し、会社自体の運営にのみ関わるタスクを実行する部門である。バック・オフィスのタスクの例としては、IT 部門、経理、人事などがある。これらのタスクは、会社の情報を処理するセキュアな E-Commerce ソフトウェアなどのバック・オフィス・システムによってサポートされる。

「[エンタープライズ・データ](#)」を参照してください。

## バッチ・プロセス

クレンジング、分類、属性の抽出、変換などを目的として、ナレッジ・ベースに大きいエンタープライズ・データ・セットを適用する操作。操作は Oracle DataLens Server で実行される。

## 標準化

標準化の目的は、データを一貫性のある明瞭で完全なものにすることである。これは、関連する製品が共通の用語と形式を使用して列記されるようにコンテンツを内部的に整合させることである。明瞭とは、コンテンツを作成した組織の外部の者が情報を理解できることである。完全とは、項目間の類似性を簡単に識別できることである。

## 品目定義

概して品目を指す、解析されたフレーズと用語のセットとは異なるものである。品目定義は1つを定義する。テクノロジーは外部の製品定義とセマンティック解析とを統合する。つまり、「トップ・ダウン」の定義と「ボトム・アップ」の解析とを統合する。

## プラットフォーム

プラットフォームとは、他の技術またはプロセスを構築する基となるソフトウェアまたはハードウェア技術の基盤である。Oracle Enterprise Data Quality for Product Data Knowledge Studio は、セマンティック情報を取得し、企業内の分散データ間のセマンティック関係に関する記録のシステムを維持するプラットフォームである。

## フレーズ構造ルール

フレーズ構造ルール (Oracle Enterprise Data Quality for Product Data Knowledge Studio の Phrase Structure フォルダで定義されている項目) は、大きい単位のナレッジ / 概念を構成する用語の順序を定義する。

## フレーズのあいまい表現

同じ用語を起源とする複数のフレーズ。

## 分類

通常は製品タイプの分類である、スキーマ内のエンタープライズ・データを識別する処理。

## 別名

属性に対して指定できるオプションの名前。スペースを含む ASCII 文字の任意の組合せを使用できる。別名は、属性名で必要とされるアンダースコアを使用せずに出力列データに名前を付ける便利な方法。

## 変換

変換の一例としては、標準化、分類、明細項目の翻訳などがある。

## 翻訳

エンタープライズ・データをある言語から別の言語に変換する Oracle Enterprise Data Quality for Product Data のプロセス。

## 翻訳品質メトリック

1 行のエンタープライズ・データの翻訳の正確さを推定するためにシステムが使用する 0 ~ 1 の範囲の値。Q 値が 1 に近いほど、翻訳を許容できる可能性が高くなる。メトリックの数は、明細項目の解析、ロケール属性、および用語集エントリの関数である。

## 翻訳用語集

ユーザーが Oracle Enterprise Data Quality for Product Data Knowledge Studio を使用して構築した言語固有の辞書。1 つのプロジェクトに対し、任意の数の翻訳辞書を作成し、コンテンツを複数の異なる言語に翻訳できる。

## 明細項目

変換が実行されるデータの単位。明細項目の例としては、最小在庫管理単位 (SKU)、製品説明、属性などがある。カスタマが提供する識別子は明細項目を一意に識別する。

## 用語ルール

用語ルールとは、同じことを意味する様々な用語および略語を参照する構造である。

## リアルタイム

ユーザーが十分に即時と感じる、またはコンピュータがなんらかの外部プロセスに遅れずに付いていくことができる (たとえば、絶えず変化する天候の視覚的表示)、コンピュータの応答性のレベル。リアルタイムで動作するコンピュータまたはプロセスに関する形容詞である。リアルタイムは、機械ではなく人間を感じる時間の説明である。

## ロケール

特定の言語およびその他の文化的慣習を使用するドメイン。

