



Informatica®
10.1

例外管理ガイド

Informatica 例外管理ガイド

10.1

2016 年 6 月

© 著作権 Informatica LLC 1998, 2018

本ソフトウェアおよびマニュアルには、Informatica LLC の所有権下にある情報が収められています。これらは使用および開示の制限等を定めた使用許諾契約のもとに提供され、著作権法により保護されています。当該ソフトウェアのリバースエンジニアリングは禁じられています。本マニュアルのいかなる部分も、いかなる手段（電子的複製、写真複製、録音など）によっても、Informatica LLC の事前の承諾なしに複製または転載することは禁じられています。このソフトウェアは、米国および/または国際的な特許、およびその他の出願中の特許によって保護されています。

合衆国政府によるソフトウェアの使用、複製または開示は、DFARS 227.7202-1 (a) および 227.7702-3 (a) (1995 年)、DFARS 252.227-7013(C) (1) (ii) (1988 年 10 月)、FAR 12.212 (a) (1995 年)、FAR 52.227-19、または FAR 52.227-14 (ALT III) に記載されているとともに、当該ソフトウェア使用許諾契約に定められた制限によって規制されます。

本製品または本書の情報は、予告なしに変更されることがあります。お客様が本製品または本書内に問題を発見された場合は、書面に当社までお知らせください。

Informatica、Informatica Platform、Informatica Data Services、PowerCenter、PowerCenterRT、PowerCenter Connect、PowerCenter Data Analyzer、PowerExchange、PowerMart、Metadata Manager、Informatica Data Quality、Informatica Data Explorer、Informatica B2B Data Transformation、Informatica B2B Data Exchange、Informatica On Demand、Informatica Identity Resolution、Informatica Application Information Lifecycle Management、Informatica Complex Event Processing、Ultra Messaging、Informatica Master Data Management、および Live Data Map は、Informatica LLC の米国および世界中の管轄地での商標または登録商標です。その他のすべての企業名および製品名は、それぞれの企業の商標または登録商標です。

本ソフトウェアまたはドキュメントの一部は、次のサードパーティが有する著作権に従います（ただし、これらに限定されません）。Copyright DataDirect Technologies. All rights reserved. Copyright (C) Sun Microsystems. All rights reserved. Copyright (C) RSA Security Inc. All rights reserved. Copyright (C) Ordinal Technology Corp. All rights reserved. Copyright (C) Aandacht c.v. All rights reserved. Copyright Genivia, Inc. All rights reserved. Copyright Isomorphic Software. All rights reserved. Copyright (C) Meta Integration Technology, Inc. All rights reserved. Copyright (C) Intalio. All rights reserved. Copyright (C) Oracle. All rights reserved. Copyright (C) Adobe Systems Incorporated. All rights reserved. Copyright (C) DataArt, Inc. All rights reserved. Copyright (C) ComponentSource. All rights reserved. Copyright (C) Microsoft Corporation. All rights reserved. Copyright (C) Rogue Wave Software, Inc. All rights reserved. Copyright (C) Teradata Corporation. All rights reserved. Copyright (C) Yahoo! Inc. All rights reserved. Copyright (C) Glyph & Cog, LLC. All rights reserved. Copyright (C) Thinkmap, Inc. All rights reserved. Copyright (C) Clearpace Software Limited. All rights reserved. Copyright (C) Information Builders, Inc. All rights reserved. Copyright (C) OSS Nokalva, Inc. All rights reserved. Copyright Edifecs, Inc. All rights reserved. Copyright Cleo Communications, Inc. All rights reserved. Copyright (C) International Organization for Standardization 1986. All rights reserved. Copyright (C) ej-technologies GmbH. All rights reserved. Copyright (C) Jaspersoft Corporation. All rights reserved. Copyright (C) International Business Machines Corporation. All rights reserved. Copyright (C) yWorks GmbH. All rights reserved. Copyright (C) Lucent Technologies. All rights reserved. Copyright (C) University of Toronto. All rights reserved. Copyright (C) Daniel Veillard. All rights reserved. Copyright (C) Unicode, Inc. Copyright IBM Corp. All rights reserved. Copyright (C) MicroQuill Software Publishing, Inc. All rights reserved. Copyright (C) PassMark Software Pty Ltd. All rights reserved. Copyright (C) LogiXML, Inc. All rights reserved. Copyright (C) 2003-2010 Lorenzi Davide, All rights reserved. Copyright (C) Red Hat, Inc. All rights reserved. Copyright (C) The Board of Trustees of the Leland Stanford Junior University. All rights reserved. Copyright (C) EMC Corporation. All rights reserved. Copyright (C) Flexera Software. All rights reserved. Copyright (C) Jinfonet Software. All rights reserved. Copyright (C) Apple Inc. All rights reserved. Copyright (C) Telerik Inc. All rights reserved. Copyright (C) BEA Systems. All rights reserved. Copyright (C) PDFlib GmbH. All rights reserved. Copyright (C) Orientation in Objects GmbH. All rights reserved. Copyright (C) Tanuki Software, Ltd. All rights reserved. Copyright (C) Ricebridge. All rights reserved. Copyright (C) Sencha, Inc. All rights reserved. Copyright (C) Scalable Systems, Inc. All rights reserved. Copyright (C) jQWidgets. All rights reserved. Copyright (C) Tableau Software, Inc. All rights reserved. Copyright (C) MaxMind, Inc. All rights reserved. Copyright (C) TMate Software s.r.o. All rights reserved. Copyright (C) MapR Technologies Inc. All rights reserved. Copyright (C) Amazon Corporate LLC. All rights reserved. Copyright (C) Highsoft. All rights reserved. Copyright (C) Python Software Foundation. All rights reserved. Copyright (C) BeOpen.com. All rights reserved. Copyright (C) CNRI. All rights reserved.

本製品には、Apache Software Foundation (<http://www.apache.org/>) によって開発されたソフトウェア、およびさまざまなバージョンの Apache License（まとめて「License」と呼んでいます）の下に許諾された他のソフトウェアが含まれます。これらのライセンスのコピーは、<http://www.apache.org/licenses/> で入手できます。適用法にて要求されないか書面に合意されない限り、ライセンスの下に配布されるソフトウェアは「現状のまま」で配布され、明示的あるいは黙示的かを問わず、いかなる種類の保証や条件も付帯することはありません。ライセンス下での許諾および制限を定める具体的文言については、ライセンスを参照してください。

本製品には、Mozilla (<http://www.mozilla.org/>) によって開発されたソフトウェア、ソフトウェア copyright The JBoss Group, LLC、コンテンツの無断複製・転載を禁じます、ソフトウェア copyright, Red Hat Middleware, LLC、コンテンツの無断複製・転載を禁じます、Copyright (C) 1999-2006 by Bruno Lowagie and Paulo Soares および GNU Lesser General Public License Agreement (<http://www.gnu.org/licenses/lgpl.html> を参照) に基づいて許諾されたその他のソフトウェアが含まれています。資料は、Informatica が無料で提供しており、一切の保証を伴わない「現状渡し」で提供されるものとし、Informatica LLC は市場性および特定の目的の適合性の黙示の保証などを含めて、一切の明示的及び黙示的保証の責任を負いません。

製品には、ワシントン大学、カリフォルニア大学アーバイン校、およびバンダービルト大学の Douglas C. Schmidt および同氏のリサーチグループが著作権を持つ ACE (TM) および TAO (TM) ソフトウェアが含まれています。Copyright (C) 1993-2006, All rights reserved.

本製品には、OpenSSL Toolkit を使用するために OpenSSL Project が開発したソフトウェア（copyright The OpenSSL Project. コンテンツの無断複製・転載を禁じます）が含まれています。また、このソフトウェアの再配布は、<http://www.openssl.org> および <http://www.openssl.org/source/license.html> にある使用条件に従います。

本製品には、Curl ソフトウェア Copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>が含まれます。All Rights Reserved. 本ソフトウェアに関する許諾および制限は、<http://curl.haxx.se/docs/copyright.html> にある使用条件に従います。すべてのコピーに上記の著作権情報とこの許諾情報が記載されている場合、目的に応じて、本ソフトウェアの使用、コピー、変更、ならびに配布が有償または無償で許可されます。

本製品には、MetaStuff, Ltd. のソフトウェアが含まれます。Copyright 2001-2005 (C) MetaStuff, Ltd. All Rights Reserved. 本ソフトウェアに関する許諾および制限は、<http://www.dom4j.org/license.html> にある使用条件に従います。

製品には、The Dojo Foundation のソフトウェアが含まれます。Copyright (C) 2004-2007. All Rights Reserved. 本ソフトウェアに関する許諾および制限は、<http://dojotoolkit.org/license> にある使用条件に従います。

本製品には、ICU ソフトウェアおよび他のソフトウェアが含まれます。Copyright International Business Machines Corporation. All rights reserved. 本ソフトウェアに関する許諾および制限は、<http://source.icu-project.org/repos/icu/icu/trunk/license.html> にある使用条件に従います。

本製品には、Per Bothner のソフトウェアが含まれます。Copyright (C) 1996-2006. All rights reserved. お客様がこのようなソフトウェアを使用するための権利は、ライセンスで規定されています。<http://www.gnu.org/software/kawa/Software-License.html> を参照してください。

本製品には、OSSP UUID ソフトウェアが含まれます。Copyright (C) 2002 Ralf S. Engelschall, Copyright (C) 2002 The OSSP Project Copyright (C) 2002 Cable & Wireless Deutschland. 本ソフトウェアに関する許諾および制限は、<http://www.opensource.org/licenses/mit-license.php> にある使用条件に従います。

本製品には、Boost (<http://www.boost.org/>) によって開発されたソフトウェア、または Boost ソフトウェアライセンスの下で開発されたソフトウェアが含まれます。本ソフトウェアに関する許諾および制限は、http://www.boost.org/LICENSE_1_0.txt にある使用条件に従います。

本製品には、University of Cambridge のが含まれます。Copyright (C) 1997-2007. 本ソフトウェアに関する許諾および制限は、<http://www.pcre.org/license.txt> にある使用条件に従います。

本製品には、The Eclipse Foundation のソフトウェアが含まれます。Copyright (C) 2007. All Rights Reserved. 本ソフトウェアに関する許諾および制限は、<http://www.eclipse.org/org/documents/epl-v10.php> および <http://www.eclipse.org/org/documents/edl-v10.php> にある使用条件に従います。

本製品には、<http://www.tcl.tk/software/tcltk/license.html>、<http://www.bosrup.com/web/overlib/?License>、<http://www.stlport.org/doc/license.html>、<http://www.asm.ow2.org/license.html>、<http://www.cryptix.org/LICENSE.TXT>、<http://hsqldb.org/web/hsqldbLicense.html>、<http://httpunit.sourceforge.net/doc/license.html>、<http://jung.sourceforge.net/license.txt>、http://www.zip.org/zlib/zlib_license.html、<http://www.openldap.org/software/release/license.html>、<http://www.libssh2.org>、<http://slf4j.org/license.html>、<http://www.sente.ch/software/OpenSourceLicense.html>、<http://fusesource.com/downloads/license-agreements/fuse-message-broker-v-5-3-license-agreement>、<http://antlr.org/license.html>、<http://aopalliance.sourceforge.net/>、<http://www.bouncycastle.org/license.html>、<http://www.jgraph.com/jgraphdownload.html>、<http://www.jcraft.com/jsch/LICENSE.txt>、http://jotm.objectweb.org/bsd_license.html に基づいて許諾されたソフトウェアが含まれています。<http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231>、<http://www.slf4j.org/license.html>、<http://nanoxml.sourceforge.net/orig/copyright.html>、<http://www.json.org/license.html>、<http://forge.ow2.org/projects/javaservice/>、<http://www.postgresql.org/about/licence.html>、<http://www.sqlite.org/copyright.html>、<http://www.tcl.tk/software/tcltk/license.html>、<http://www.jaxen.org/faq.html>、<http://www.jdom.org/docs/faq.html>、<http://www.slf4j.org/license.html>、<http://www.iodbc.org/dataspace/iodbc/wiki/iODBC/License>、<http://www.keplerproject.org/md5/license.html>、<http://www.toedter.com/en/jcalendar/license.html>、<http://www.edankert.com/bounce/index.html>、<http://www.net-snmp.org/about/license.html>、<http://www.openmdx.org/#FAQ>、http://www.php.net/license/3_01.txt、<http://srp.stanford.edu/license.txt>、<http://www.schneider.com/blowfish.html>、<http://www.jmock.org/license.html>、<http://xsom.java.net>、<http://benalman.com/about/license/>、<https://github.com/CreateJS/EaselJS/blob/master/src/easeljs/display/Bitmap.js>、<http://www.h2database.com/html/license.html#summary>、<http://jsoncpp.sourceforge.net/LICENSE>、<http://jdbc.postgresql.org/license.html>、<http://protobuf.googlecode.com/svn/trunk/src/google/protobuf/descriptor.proto>、<https://github.com/rantav/hector/blob/master/LICENSE>、<http://web.mit.edu/Kerberos/krb5-current/doc/mitK5license.html>、<http://jibx.sourceforge.net/jibx-license.html>、<https://github.com/lyokato/libgeohash/blob/master/LICENSE>、<https://github.com/hjiang/jsonxx/blob/master/LICENSE>、<https://code.google.com/p/lz4/>、<https://github.com/jedisct1/libsodium/blob/master/LICENSE>、<http://one-jar.sourceforge.net/index.php?page=documents&file=license>、<https://github.com/EsotericSoftware/kryo/blob/master/license.txt>、<http://www.scala-lang.org/license.html>、<https://github.com/tinkerpop/blueprints/blob/master/LICENSE.txt>、<http://gee.cs.oswego.edu/dl/classes/EDU/oswego/cs/dl/util/concurrent/intro.html>、<https://aws.amazon.com/asl/>、<https://github.com/twbs/bootstrap/blob/master/LICENSE>、および <https://sourceforge.net/p/xmlunit/code/HEAD/tree/trunk/LICENSE.txt>。

本製品には、Academic Free License (<http://www.opensource.org/licenses/afl-3.0.php>)、Common Development and Distribution License (<http://www.opensource.org/licenses/cddl1.php>)、Common Public License (<http://www.opensource.org/licenses/cpl1.0.php>)、Sun Binary Code License Agreement Supplemental License Terms、BSD License (<http://www.opensource.org/licenses/bsd-license.php>)、BSD License (<http://opensource.org/licenses/BSD-3-Clause>)、MIT License (<http://www.opensource.org/licenses/mit-license.php>)、Artistic License (<http://www.opensource.org/licenses/artistic-license-1.0>)、Initial Developer's Public License Version 1.0 (<http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/>) に基づいて許諾されたソフトウェアが含まれています。

本製品には、ソフトウェア copyright (C) 2003-2006 Joe Walnes, 2006-2007 XStream Committers が含まれています。All rights reserved.本ソフトウェアに関する許諾および制限は、<http://j.org/license.html> にある使用条件に従います。本製品には、Indiana University Extreme! Lab によって開発されたソフトウェアが含まれています。詳細については、<http://www.extreme.indiana.edu/> を参照してください。

本製品には、ソフトウェア Copyright (C) 2013 Frank Balluffi and Markus Moeller が含まれています。All rights reserved.本ソフトウェアに関する許諾および制限は、MIT ライセンスの使用条件に従います。

特許については、<https://www.informatica.com/legal/patents.html> を参照してください。

免責: 本文書は、一切の保証を伴わない「現状渡し」で提供されるものとし、Informatica LLC は他社の権利の非侵害、市場性および特定の目的への適合性の黙示の保証などを含めて、一切の明示的および黙示的保証の責任を負いません。Informatica LLC では、本ソフトウェアまたはドキュメントに誤りのないことを保証していません。本ソフトウェアまたはドキュメントに記載されている情報には、技術的に不正確な記述や誤植が含まれる場合があります。本ソフトウェアまたはドキュメントの情報は、予告なしに変更されることがあります。

NOTICES

この Informatica 製品（以下「ソフトウェア」）には、Progress Software Corporation（以下「DataDirect」）の事業子会社である DataDirect Technologies からの特定のドライバ（以下「DataDirect ドライバ」）が含まれています。DataDirect ドライバには、次の用語および条件が適用されます。

1. DataDirect ドライバは、特定物として現存するままの状態提供され、商品性の保証、特定目的適合性の保証および法律上の瑕疵担保責任を含むすべての明示もしくは黙示の保証責任を負わないものとします。国または地域によっては、法律の強行規定により、保証責任の制限が禁じられる場合、強行規定の制限を受けるものとします。
2. DataDirect または第三者は、予見の有無を問わず発生した ODBC ドライバの使用に関するいかなる直接的、間接的、偶発的、特別、あるいは結果的損害に対して責任を負わないものとします。本制限事項は、すべての訴訟原因に適用されます。訴訟原因には、契約違反、保証違反、過失、厳格責任、詐称、その他の不法行為を含みますが、これらに限るものではありません。

発行日: 2018-07-04

目次

序文	7
Informatica のリソース	7
Informatica Network	7
Informatica ナレッジベース	7
Informatica マニュアル	8
Informatica 製品可用性マトリックス	8
Informatica Velocity	8
Informatica Marketplace	8
Informatica グローバルカスタマサポート	8
第 1 章 : 例外管理の概要	9
例外管理の概要	9
タスクのタイプ	9
タスクとワークフロー	10
ヒューマンタスクのステップ	11
例外管理におけるロール	11
[例外管理] ワークスペース	12
[開始] ワークスペース	12
[例外] ワークスペース	14
監査証跡の操作	14
例外管理プロセスフロー	14
不良レコードの例	15
重複レコードの例	16
Analyst ツールへのログイン	17
第 2 章 : 例外タスク	18
例外タスクの概要	18
例外タスクのタイプ	18
例外修正の手順	19
例外確認の手順	19
[例外] ワークスペースの例外タスク	20
例外タスクのエラーの修正	21
例外タスクの検証エラー	22
例外レコードのステータスインジケータ	22
例外レコードのステータスの設定	23
例外タスクのフィルタオプション	24
レコードデータフィルタ	25
[データ編集] タブのメタデータフィルタ	25
[データ監査] タブのメタデータフィルタ	26
例外タスクのレコードのフィルタリング	26

例外レコードでの値の検索および置換.	27
例外レコードでの単一値の置換.	27
例外レコードでの複数の値の置換.	28
レコードへのメモの追加.	29
確認中の例外タスク.	29
例外確認タスクでの作業.	29
例外タスクのルールおよびガイドライン.	30
第3章: 重複レコードタスク.	31
重複レコードタスクの概要.	31
優先レコード.	32
重複レコードタスクのタイプ.	32
重複の修正手順.	32
重複の確認手順.	33
[例外] ワークスペースの重複レコードタスク.	33
クラスタの編集.	35
重複レコードのステータスインジケータ.	36
クラスタのステータスの設定.	37
複数のクラスタでのレコードの検索.	37
クラスタ間でのレコードの移動.	38
空のクラスタへのレコードの移動.	39
クラスタへのメモの追加.	39
重複タスクのフィルタオプション.	40
重複レコードタスクでのデータフィルタ.	40
重複レコードタスクでの監査証跡フィルタ.	40
重複レコードタスクでのクラスタのフィルタリング.	41
重複確認タスク.	41
重複確認タスクでの作業.	42
重複レコードタスクのルールおよびガイドライン.	43
第4章: タスク管理.	44
タスク管理の概要.	44
タスク実行者の操作.	44
タスクの表示.	45
タスクを開く.	45
タスクの解放.	46
ビジネス管理者の操作.	46
ユーザーへのタスクの割り当て.	47
ヒューマンタスクでのタスクインスタンスのリストの表示.	47
タスクサマリデータ.	48
複数タスクの完了.	48
タスクデータのエクスポート.	49
例外タスクのメタデータ.	49

重複タスクのメタデータ.....	50
タスクデータのエクスポート.....	51
索引.....	52

序文

『*Informatica Exception Management ガイド*』では、Analyst ツールでの Exception Management の使用方法について説明します。

Exception Management は、データ品質の例外レコードをヒューマンタスクで表示および更新できる、Analyst ツールの機能です。例外とは、不良データまたは重複データを含む可能性のあるレコードのことです。Exception Management を使用すると、データエラーを解決し、重複レコードのクラスタを 1 つのレコードに統合できます。

Informatica のリソース

Informatica Network

Informatica Network は、Informatica グローバルカスタマサポート、Informatica ナレッジベースなどの製品リソースをホストします。Informatica Network には、<https://network.informatica.com> からアクセスしてください。

メンバーは以下の操作を行うことができます。

- 1 つの場所からすべての Informatica のリソースにアクセスできます。
- ドキュメント、FAQ、ベストプラクティスなどの製品リソースをナレッジベースで検索できます。
- 製品の提供情報を表示できます。
- 自分のサポート事例を確認できます。
- 最寄りの Informatica ユーザーグループネットワークを検索して、他のユーザーと共同作業を行えます。

Informatica ナレッジベース

ドキュメント、ハウツー記事、ベストプラクティス、PAM などの製品リソースを Informatica Network で検索するには、Informatica ナレッジベースを使用します。

ナレッジベースには、<https://kb.informatica.com> からアクセスしてください。ナレッジベースに関する質問、コメント、ご意見の連絡先は、Informatica ナレッジベースチーム (KB_Feedback@informatica.com) です。

Informatica マニュアル

使用している製品の最新のドキュメントを取得するには、
https://kb.informatica.com/_layouts/ProductDocumentation/Page/ProductDocumentSearch.aspx にあ
る Informatica ナレッジベースを参照してください。

このマニュアルに関する質問、コメント、ご意見の電子メールの送付先は、Informatica マニュアルチーム
(infa_documentation@informatica.com) です。

Informatica 製品可用性マトリックス

製品可用性マトリックス（PAM）には、製品リリースでサポートされるオペレーティングシステム、データベ
ースなどのデータソースおよびターゲットが示されています。Informatica Network メンバである場合は、
PAM
(<https://network.informatica.com/community/informatica-network/product-availability-matrices>) に
アクセスできます。

Informatica Velocity

Informatica Velocity は、Informatica プロフェッショナルサービスによって開発されたヒントおよびベスト
プラクティスのコレクションです。数多くのデータ管理プロジェクトの経験から開発された Informatica
Velocity には、世界中の組織と協力して優れたデータ管理ソリューションの計画、開発、展開、および維持を
行ってきた弊社コンサルタントの知識が集約されています。

Informatica Network メンバである場合は、Informatica Velocity リソース
(<http://velocity.informatica.com>) にアクセスできます。

Informatica Velocity についての質問、コメント、またはアイデアがある場合は、ips@informatica.com から
Informatica プロフェッショナルサービスにお問い合わせください。

Informatica Marketplace

Informatica Marketplace は、お使いの Informatica 製品を強化したり拡張したりするソリューションを検索
できるフォーラムです。Informatica の開発者およびパートナーの何百ものソリューションを利用して、プロ
ジェクトで実装にかかる時間を短縮したり、生産性を向上させたりできます。Informatica Marketplace には、
<https://marketplace.informatica.com> からアクセスできます。

Informatica グローバルカスタマサポート

Informatica Network の電話またはオンラインサポートからグローバルカスタマサポートに連絡できます。

各地域の Informatica グローバルカスタマサポートの電話番号は、Informatica Web サイト
(<http://www.informatica.com/us/services-and-training/support-services/global-support-centers>) を参
照してください。

Informatica Network メンバである場合は、オンラインサポート (<http://network.informatica.com>) を使用
できます。

第 1 章

例外管理の概要

この章では、以下の項目について説明します。

- [例外管理の概要, 9 ページ](#)
- [タスクのタイプ, 9 ページ](#)
- [タスクとワークフロー, 10 ページ](#)
- [例外管理におけるロール, 11 ページ](#)
- [\[例外管理\] ワークスペース, 12 ページ](#)
- [監査証跡の操作, 14 ページ](#)
- [例外管理プロセスフロー, 14 ページ](#)
- [Analyst ツールへのログイン, 17 ページ](#)

例外管理の概要

例外は、未解決のデータ品質問題を含むレコードです。このようなレコードにはエラー、または冗長な情報や重複する情報が含まれている可能性があります。例外レコードは、Analyst ツールで確認および修正できます。

例外レコードに対する作業を実行するには、タスクを開きます。タスクは、データベーステーブル内の例外レコードのセットを識別します。

Analyst ツールの「開始」ワークスペースに、作業が可能なタスクが表示されます。「開始」ワークスペースには、同じデータベーステーブルまたは異なるデータベーステーブルのタスクが含まれています。このワークスペースからタスクを開いたユーザーがタスクの所有権を要求し、タスクに指定されたデータ品質の問題の解決を担当します。

タスクのタイプ

ワークフローによってユーザーに割り当てられるタスクのタイプは、ソースデータベーステーブルに含まれているデータ品質の問題によって異なります。テーブル内の各レコードには、エラー、NULL 値、または現在のデータプロジェクトでは不正確な値が格納されていることがあります。テーブルには冗長なレコード（内容は同じだがバージョンのみが異なる情報）が格納されていることがあります。

Analyst ツールでは、次のタイプのタスクインスタンスを実行できます。

例外修正タスク

エラーまたは NULL 値が含まれている可能性のあるレコードが格納されています。見つかったエラーを分析および修正します。レコードにエラーがなくなったら、レコードが有効であることを示すようにレコードのステータスを更新します。

重複修正タスク

重複する情報が含まれている可能性のあるレコードが格納されています。このタスクはクラスタ内のレコードをソートします。クラスタとは、ソースデータセット内の同じビジネスエンティティを表すレコードのグループのことです。各クラスタを分析して、クラスタによって表される優先バージョンのレコードを定義します。優先レコードを確認したことを示すようクラスタのステータスを更新します。

レコードがクラスタ内のどのレコードとも重複していない場合は、そのレコードを他のクラスタに移動します。単一の一意のレコードを含むクラスタを作成することができます。

例外確認タスク

前のユーザーが例外修正タスクで分析した例外レコードのセットが格納されています。前のユーザーによって実行された作業を確認し、レコードデータとレコードステータスが正しいことを確認します。

重複確認タスク

前のユーザーが重複修正タスクで実行したクラスタのセットが格納されています。前のユーザーによって実行された作業を確認し、優先レコードデータとクラスタステータスが正しいことを確認します。

注: 例外レコードを修正または確認するタスクから、重複レコードを修正または確認するタスクに、レコードを引き渡すことはできません。例外レコードと重複レコードとは、格納されるデータベーステーブルの構造が異なります。

タスクとワークフロー

ワークフローと呼ばれる Informatica プロセスにより、Analyst ツールでユーザーが作業するタスクが生成されます。

ワークフローは、Informatica アプリケーションで実行できるタスクのシーケンスを記述するものです。タスクには、マッピングを実行するための指示や、電子メールを送信するための指示、Analyst ツールユーザーに例外レコードを配布するための指示が含まれる場合があります。Analyst ツールユーザーに例外レコードを配布するタスクをヒューマンタスクと呼びます。Analyst ツールで作業するタスクはヒューマンタスクインスタンスです。各タスクインスタンスで、例外データベーステーブルの一連のレコードが特定されます。

ヒューマンタスクを含むワークフローには、マッピングタスクも含まれている必要があります。マッピングタスクでは、マッピングと呼ばれるデータ処理が実行されます。マッピングによってソースデータのレコードの品質が確認され、一連のデータ品質インジケータに基づいてさまざまなテーブルにレコードが書き込まれます。ヒューマンタスクでは、例外レコードを含むテーブルが読み取られて、そのレコードが Analyst ツールユーザーに割り当てられ、レコードの確認および修正が行われます。

例外レコードテーブルに大量のレコードが含まれる場合、タスクインスタンスでレコードのサブセットが特定されます。タスクインスタンスは、ワークフローにより、ワークフロー開発者が特定した Analyst ツールユーザーに配布されます。Analyst ツールユーザーは、各タスクインスタンスの作業を同時に進めることができます。

ヒューマンタスクのステップ

ワークフローのヒューマンタスクには1つ以上のステップが含まれます。ステップは、ヒューマンタスクに流れるデータのフローを定義するメタデータオブジェクトです。ヒューマンタスクで作成され、作業が可能なユーザーが識別されるタスクインスタンスのタイプは、ステップにより決まります。

ワークフロー開発者はヒューマンタスクに次のステップを追加できます。

例外ステップ

例外ステップは Analyst ツールで例外修正タスクインスタンスを生成します。

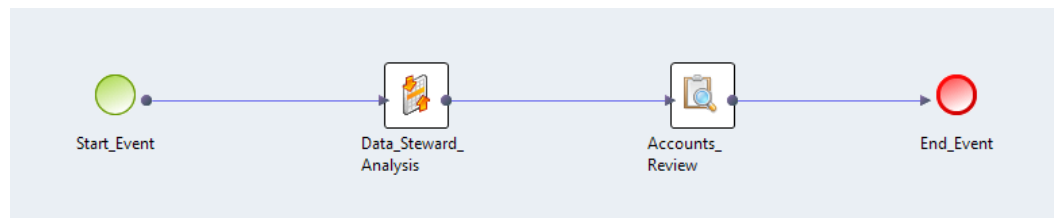
クラスタステップ

クラスタステップは Analyst ツールで重複修正タスクインスタンスを生成します。

確認ステップ

確認ステップは Analyst ツールで例外確認タスクインスタンスまたは重複確認タスクインスタンスを生成します。

次の図に、Developer tool のヒューマンタスクのステップを示します。



ヒューマンタスクに複数のステップが含まれる場合、例外レコードが1つのステップから別のステップに渡されます。ワークフローを設定する開発者はステップのシーケンスを定義します。各ステップは Analyst ツールユーザーのセットを特定します。

例外管理におけるロール

ヒューマンタスクでは、タスクインスタンスを処理する Analyst ツールユーザーにロールが割り当てられます。ロールにより、ユーザーが実行できるアクションの範囲が決まります。ヒューマンタスクでは、ロールをユーザーまたはグループに割り当てることができます。グループに属しているユーザーは、Analyst ツールにログインするとグループロールを共有できます。

ヒューマンタスクでは次のロールが割り当てられます。

タスクの実行者

Analyst ツールでタスクインスタンスを処理するユーザーまたはグループ。

ビジネス管理者

ヒューマンタスクで生成されるタスクインスタンスを管理するユーザーまたはグループ。ビジネス管理者もタスクインスタンスを処理することができます。

ヒューマンタスクでタスクインスタンスがグループに割り当てられている場合、そのグループのどのメンバーもタスクの所有権を取得できます。1人のグループメンバーがタスクの所有権を取得すると、そのタスクは他のグループメンバーのタスクリストからなくなります。

ワークフロー開発者は、ヒューマンタスクのステップごとに1人以上のタスク実行者を特定します。ユーザーはヒューマンタスクの1つのステップのタスクインスタンスのみを処理し、他のステップのタスクインスタンスを処理しない場合があります。

ワークフロー開発者は、ヒューマンタスクに対して1人以上のビジネス管理者を特定します。ヒューマンタスクのビジネス管理者は、ヒューマンタスクの各ステップのビジネス管理者でもあります。

[例外管理] ワークスペース

Analyst ツールでは、ワークスペースを使用して、ユーザーが実行できるさまざまなタイプのデータ操作が整理されます。

[開始] ワークスペースの【マイタスク】ビューに、ユーザーが所有するタスクのリストが表示されます。タスクを開くと、【例外】ワークスペースにタスクレコードが表示されます。

[開始] ワークスペース

[開始] ワークスペースには、所有しているタスクと管理しているすべてのタスクが表示されます。[開始] ワークスペースの各オプションを使用して、タスクの開始、タスクに対するアクションの実行、タスクメタデータの確認を行います。

Analyst ツールへのログインで使用するユーザーロールによって、[開始] ワークスペースに表示できるタスクが決まります。【マイタスク】ビューを選択すると、所有しているタスクが表示されます。【タスク管理】ビューを選択すると、管理しているすべてのタスクが表示されます。タスクを作成したワークフローでユーザーがビジネス管理者として識別された場合、ユーザーはタスクを管理できます。

次の画像は、【マイタスク】ビューのタスクのリストを示しています。

ID	Task Title	Task Type	Due Date	Status	Owner	Created
25	cleanse_a {1 - 8}	Correct Exceptions	09/17/2015	Assigned	docs	07/29/2015
26	cleanse_a {9 - 16}	Correct Exceptions	09/17/2015	Assigned	docs	07/29/2015
33	cleanse_a {17 - 24}	Correct Exceptions	09/17/2015	Assigned	docs	07/29/2015

ID:	25
Task Title:	cleanse_a {1 - 8}
Task Type:	Correct Exceptions
Due Date:	09/17/2015 05:01:13 PM
Status:	Assigned
Owner:	docs
Created:	07/29/2015 05:01:13 PM
Comments:	No Comments

【マイタスク】には次のオプションがあります。

1. 【マイタスク】および【タスク管理】オプション。所有しているタスクと管理しているすべてのタスク間でワークスペースビューを切り替えます。【マイタスク】ビューと【タスク管理】ビューからタスクを開くことができます。
2. 更新。ワークスペースのデータを更新します。
3. 【アクション】メニュー。選択したタスクに対して実行できるアクションのリストを開きます。
4. カラムの見出し。タスクインスタンスを説明するカラムの名前を表示します。
5. フィルタ。入力した値を使用してレコードのリストをフィルタリングします。
6. タスク名。タスクの名前を表示します。タスクを開くには、タスク名を選択します。

7. コメント。使用できるエディタを開いてコメントを入力したり、タスクに関するコメントを確認したりします。

[開始] ワークスペースのカラムの見出し

[開始] ワークスペースのテーブルに各タスクが表示されます。各タスクはテーブルに 1 行で表示されます。テーブルのカラムには、タスクのメタデータが含まれています。

次の各カラムを参照して、自分が所有または管理しているタスクのメタデータを確認できます。

ID

タスクインスタンスの一意の ID。ID には、タスクインスタンスを生成したヒューマンタスクを識別する値が含まれています。

タスクのタイトル

タスクの名前。タスク名は、タスクデータを生成したヒューマンタスクの手順を示します。

タスクタイプ

タスクデータに対して実行できるデータ操作のタイプ。

タスクのタイプは次のいずれかです。

- 例外修正。レコードのエラーを調べて修正します。
- 重複修正。重複レコードのクラスタを調べて、レコードの値から優先レコードを作成します。
- 例外確認。タスク内の別のユーザーによる例外修正作業を確認します。
- 重複確認。タスク内の別のユーザーによるレコードクラスタからの優先レコードの作成作業を確認します。

注: タスクタイプには投票タスクも含まれることがあります。Business Glossary ユーザーが投票タスクを実行します。投票タスクは、例外管理プロセスには含まれません。

期限

タスクの期限。ヒューマンタスクは、例外修正タスクと重複修正タスクの期限を定めます。Analyst ツールは、例外修正タスクと重複修正タスクの期限を計算します。

ステータス

タスクのステータス。

タスクのステータスは次のいずれかです。

- 作成日。タスクの所有者がいません。
- 割り当て済み。タスクの所有者がいます。

Owner (所有者)

現在のタスク所有者の名前。

作成日

ワークフローがタスクを作成した日付。

[例外] ワークスペース

[例外] ワークスペースは、タスクを表示または開く場合に表示される一時ワークスペースです。 [例外] ワークスペースには、**[データ編集]** タブと **[データ監査]** タブがあります。

[データ編集] タブには、タスクのレコードまたはクラスタの更新に使用できるタスクデータとオプションが表示されます。 タブには、Analyst ツールでユーザーが各レコードやクラスタに対して行う更新を追跡するために使用するメタデータカラムも表示されます。

[データ監査] タブには、ユーザーがタスクデータに加えた変更の監査証跡が表示されます。 ユーザーが変更したフィールド、各レコードを変更したユーザの ID、および変更日を表示できます。

タスクに対する作業が完了したら、[例外] ワークスペースを閉じることができます。

監査証跡の操作

Analyst ツールには、タスクで行った変更の監査証跡が保存されます。 タスクの監査証跡を表示するには、[例外] ワークスペースで **[データ監査]** タブを選択します。

監査証跡には、ユーザーがタスクで更新した各レコードについて、レコードデータとステータスインジケータの値が表示されます。 また、監査証跡を確認することで、更新を実行したユーザーと更新日時も特定できます。 監査証跡では行が古い順に表示されます。

監査証跡の各行は、レコードに対する 1 つの更新を表します。 1 つのレコードに複数の更新を行った場合、監査証跡では更新ごとに行が追加されます。

レコードのデータ値を変更すると、監査証跡ではその値がハイライト表示されます。 その値にマウスポインタを移動すると、直前の値が表示されます。 監査証跡では、レコードを変更したユーザー、変更日、変更のタイプも記録されます。

監査証跡には、重複レコードタスクのクラスタメタデータは表示されません。 タスクがまったく更新されていない場合、監査証跡は空になります。

例外管理プロセスフロー

Analyst ツールで作業するタスクは、データ品質サイクルの 1 つのステージを表します。 組織がデータセットのデータ品質を確認することを決定した時点でこのサイクルは始まります。 組織がデータの品質に満足した時点でこのサイクルは終了します。 組織はデータ品質サイクルを継続的に実行することができます。

多くの場合、例外管理のステージはデータ品質サイクルの終了まで続けられます。 初期のステージでは、組織のデータ品質の測定と向上を目的として、プロファイルとマッピングを使用することもあります。 例外管理では、現在のサイクルのデータ品質ターゲットを満たしていないレコードに対してユーザーが実行できる手動操作を定義します。

ユーザーは、データ品質プロセスを実行するデータスチュワードチームの一員として作業する場合があります。 また、ユーザーは、データが準拠する必要があるデータ品質標準を定義するビジネスユーザーの場合もあります。 どちらの場合も、ユーザーは Analyst ツールで例外管理タスクを所有することができます。

注: 多くの組織で、データスチュワードはデータスチュワードシップのロールを他のロールと組み合わせます。 ユーザーは組織の中のより大きなロールの一部としてデータ品質目標に対する責任を担う場合もあれば、データスチュワードとして、データセットに対してタスクを実行し、その結果をデータスチュワードシップの責務を負う同僚に渡す場合もあります。

不良レコードの例

あなたは小売企業のデータスチュワードです。製品在庫レコードにエラーがあるのではないかと懸念しています。そのエラーが原因で企業は製品を販売能力より過大または過小に注文している可能性があります。

次のプロセスを定義してエラーを調査し修正します。

1. 開発者に依頼して、データセット内のエラーを見つけて修正する 1 つ以上のマッピングを設定してもらいます。

このマッピングでは、データセット内の各レコードの数値スコアも計算します。スコアはレコードのデータ品質を表します。中には、スコアが境界値付近であるため、マッピングでレコードに含まれるデータ品質問題の一部を検証できないレコードもあります。

2. 開発者は、数値スコアを読み取る追加のマッピングを設定します。そして、このマッピングを、マッピングタスクとヒューマンタスクを含むワークフローに追加します。

- マッピングタスクはマッピングを実行します。マッピングタスクは、レコードに格納されているスコアに基づいてレコードを別々のテーブルに書き込みます。
- ヒューマンタスクは、境界スコアを持つレコードを複数のタスクに配布します。これらのタスクは、あなたや他のユーザーが Analyst ツールで開くことができます。

3. あなたは Analyst ツールにログインして、タスクを開きます。タスクは 1 つ以上のテーブルの例外レコードをまとめます。各テーブルには 100 のレコードを含むことが可能です。

各レコードに対して次のいずれかのアクションを実行します。

- レコードのエラーを修正するか、または現在のレコードが正しいと判断する。

レコードステータスを更新して、レコードが有効であることを示します。

- レコードに有効なデータが含まれないと判断する。

レコードステータスを更新して、レコードが無効であることを示します。

- レコードの正確性を検証できないと決定する。

レコードステータスを更新して、他のユーザーまたは他の Informatica プロセスによりレコードをさらに分析する必要があることを示します。

注: レコードを更新する前に、タスクが編集モードで開いていることを確認します。編集モードに入るには、開いているタスクで **[編集]** ボタンをクリックします。

4. タスクのすべてのレコードに対する作業が完了したら、タスクのステータスを更新します。タスクのステータスは、レコードがデータ品質処理の次のステージに進める状態になったことを示します。

データの次のステージは、ヒューマンタスクの設定によって異なります。例えば、ヒューマンタスクにはレコードを確認のために他のユーザーに割り当てる追加ステップが含まれることがあります。

ヒューマンタスクが完了すると、ワークフローの次のステージが開始されます。

重複レコードの例

あなたは銀行のデータスチュワードです。顧客口座テーブルの複数のレコードに同じ情報が格納されている可能性があるのではないかと懸念しています。こうした重複するレコードは、データ入力エラーまたは顧客による不正行為の可能性を示唆しています。

そこで、次のプロセスを定義して重複レコードを検索し、各レコードセットについて1つの優先バージョンを特定することにします。

1. 開発者に依頼して、重複レコードを特定する1つ以上のマッピングを設定してもらいます。

このマッピングでは、各レコードのデータ値の重複レベルを表す数値スコアのセットを計算します。スコアが高い場合はレコードが重複していることを示し、スコアが低い場合はレコードが一意であることを示しています。中には、スコアが境界値付近であるため、レコードの重複ステータスがはっきりしないレコードもあります。

2. 開発者は、数値スコアを読み取る追加のマッピングを設定します。そして、このマッピングを、マッピングタスクとヒューマンタスクを含むワークフローに追加します。

- マッピングタスクはマッピングを実行します。マッピングタスクは、レコードに格納されているスコアに基づいてレコードを別々のテーブルに書き込みます。
- ヒューマンタスクは、境界スコアを持つレコードを複数のタスクに配布します。これらのタスクは、あなたや他のユーザーがAnalyst ツールで開くことができます。

3. あなたはAnalyst ツールにログインして、タスクを開きます。

Analyst ツールによって、レコードが一連のクラスタに編成されます。各クラスタは、同じような情報が格納されている2つ以上のレコードで構成されます。デフォルトでは、クラスタの先頭のレコードが優先レコードになっています。

4. クラスタを開き、格納されているレコードを分析します。

各クラスタについて、次のアクションを実行します。

- レコードデータの各カラムのデータ値を調べます。各カラムで最も正確な値を選択し、その値を優先レコードに昇格させます。
選択した値を編集することや、共通の値を持つレコードを他のクラスタ内で検索することができます。
- 現在のクラスタに属さないレコードは別のクラスタに移動するか、そのレコード用に新しいクラスタを作成します。
- クラスタを確認したことを示すようクラスタのステータスを更新します。各クラスタで現在の優先レコードを確認したらタスクを完了します。

注: レコードを更新する前に、タスクが編集モードで開いていることを確認します。編集モードに入るには、開いているタスクで**【編集】** ボタンをクリックします。

5. タスク内のすべてのクラスタに対する作業が完了したら、タスクのステータスを更新します。タスクのステータスは、レコードがデータ品質処理の次のステージに進める状態になったことを示します。

データの次のステージは、ヒューマンタスクの設定によって異なります。例えば、ヒューマンタスクにはクラスタを他のユーザーに確認のため割り当てる追加ステップが含まれることがあります。

ヒューマンタスクが完了すると、ワークフローの次のステージが開始されます。

Analyst ツールへのログイン

Analyst ツールにログインするには、Analyst ツールの URL を使用します。Analyst ツールにログインするときは、Informatica ログイン名、パスワード、およびネイティブドメインか LDAP セキュリティドメインを指定します。

1. Microsoft Internet Explorer または Google Chrome ブラウザを起動します。
2. [アドレス] フィールドに、次に示す Analyst ツールの URL を入力します。
`http[s]://<fully qualified host name>:<port number>/analyst/`
3. ドメインで LDAP またはネイティブ認証を使用する場合は、ログインページでログイン名とパスワードを入力します。
4. **【ネイティブ】** または特定のセキュリティドメインの名前を選択します。
Informatica ドメインで LDAP または Kerberos 認証を使用している場合は、[セキュリティドメイン] フィールドが表示されます。ユーザーアカウントが所属するセキュリティドメインがわからない場合は、Informatica ドメイン管理者に問い合わせてください。
5. **【ログイン】** をクリックします。
Analyst ツールにより、**【開始】** ワークスペースが開きます。

第 2 章

例外タスク

この章では、以下の項目について説明します。

- [例外タスクの概要, 18 ページ](#)
- [例外タスクのタイプ, 18 ページ](#)
- [\[例外\] ワークスペースの例外タスク, 20 ページ](#)
- [例外タスクのエラーの修正, 21 ページ](#)
- [例外レコードのステータスインジケータ, 22 ページ](#)
- [例外タスクのフィルタオプション, 24 ページ](#)
- [例外レコードでの値の検索および置換, 27 ページ](#)
- [レコードへのメモの追加, 29 ページ](#)
- [確認中の例外タスク, 29 ページ](#)
- [例外タスクのルールおよびガイドライン, 30 ページ](#)

例外タスクの概要

例外タスクはデータエラーを含む可能性のあるレコードを識別します。レコードには不正確なデータや NULL フィールドが含まれることがあります。ワークフローでは、前のプロセスでデータセットのデータ品質問題の一部が解決できなかった場合、例外タスクが生成されます。例外タスクを開くと、Analyst ツールで未解決の問題を含む各レコードのフィールドが識別されます。

例外タスクのタイプ

ワークフローにより、例外修正タスクまたは例外確認タスクを割り当てることができます。

例外修正タスクで作業を行う場合、タスクの各レコードを調べて、レコードデータを更新するかどうかを決定します。例外確認タスクで作業する場合、前のタスクで各レコードに対して他のユーザーが行った決定を検証します。

各タスクタイプのワークスペースレイアウトは同じです。両タイプのタスクで同じ方法によりレコードデータとステータスインジケータを更新します。例外確認タスクを行う場合、追加オプションを使用して前のユーザーの作業を承認または拒否できます。

例外修正の手順

〔開始〕 ワークスペースから例外修正タスクを開きます。〔例外〕 ワークスペースの〔データ編集〕 タブでタスクが開きます。

例外修正タスクで作業する場合、次の手順を実行します。

1. タスクを開きます。
2. 各レコードで強調表示されたデータフィールドを調べます。
矢印を使用してタスク内のレコード間を移動できます。フィルタオプションを使用してデータ品質問題を含むレコードまたは指定したステータスを表示または非表示にします。
3. データフィールドに含まれているエラーを修正します。
4. 現在のレコードデータが反映されるよう各レコードのステータスを更新します。
次のいずれかのオプションを選択します。
 - 承認。現在のデータはビジネスで承認可能と決定します。
 - 拒否。現在のデータはビジネスで承認不可と決定します。
 - 再処理。現在のデータのステータスを決定できません。
5. レコードを更新してレコードデータが確認済みであることを示します。他のステータス値とは独立して〔確認済み〕ステータス値を設定できます。
フィルタオプションを使用して共通のステータス値を持つレコードを表示または非表示にします。
6. 必要に応じて、メモをレコードに追加します。例えば、レコードを拒否した理由をメモとして追加する場合があります。
7. タスクのすべてのレコードに対する作業が完了したら、タスクのステータスを更新します。

例外確認の手順

〔開始〕 ワークスペースから例外確認タスクを開きます。〔例外〕 ワークスペースの〔データ編集〕 タブでタスクが開きます。

例外確認タスクの作業を行うときは、次の手順を実行します。

1. タスクを開きます。
2. 各レコードを確認します。レコードデータを調べ、前のユーザーが各レコードに設定したステータスインジケータを調べます。
 - レコードの現在のコンテンツに同意する場合、変更は加えません。コンテンツに同意しないレコードがあれば、レコードを更新します。
 - 現在のレコードステータスに同意する場合、変更は加えません。レコードステータスに同意しない場合、ステータスを更新します。フィルタオプションを使用して共通のステータス値を持つレコードを表示または非表示にします。
3. 各レコードの確認ステータスを検証します。確認ステータスは、レコードを承認または拒否することを示します。確認ステータスは前のユーザーがレコードに対して適用したステータスより優先されます。
4. 必要に応じて、メモをレコードに追加します。前のタスクでユーザーがレコードにメモを追加した場合、過去のメモは今回追加するメモで置換されます。
5. タスクのすべてのレコードに対する作業が完了したら、タスクのステータスを更新します。

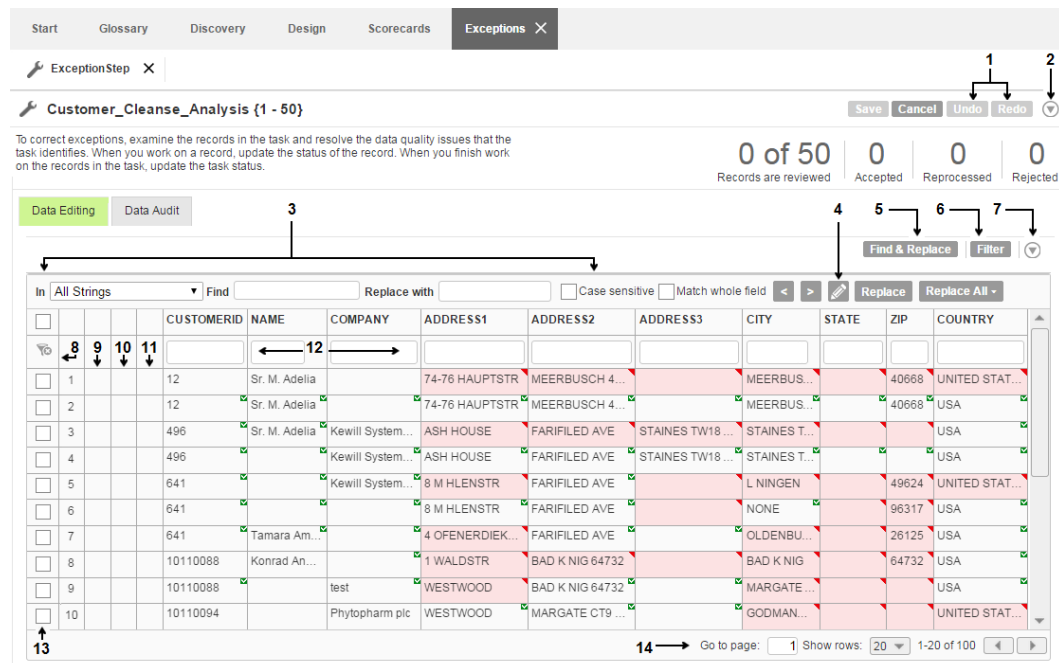
[例外] ワークスペースの例外タスク

例外タスクを開くと、[例外] ワークスペースにタスク内のレコードがリストされ、タスクに設定可能なオプションが表示されます。タスクは「データ編集」タブに表示されます。タブには例外修正タスクおよび例外確認タスクに共通のオプションセットが表示されます。

Analyst ツールは、赤い背景を使用して、未解決のデータ品質の問題が含まれるデータフィールドを強調表示します。レコードを調べて、レコード内のエラーと識別したフィールドを更新します。データ値を更新すると、Analyst ツールはセルの背景色をクリアして、そのセルに緑色のチェックマークを付けます。カーソルを使用すると、そのセルの以前の値を表示できます。

レコードに対する作業が完了したら、レコードステータスを更新します。タスク内のすべてのレコードに対する作業が完了したら、タスクステータスを更新してレコードがワークフローの次のステージに進める状態になったことを示します。

次の画像は、例外レコードが含まれるタスクを編集モードで開いたときの **【データ編集】** タブを示しています。



タスクに対する作業を行うときは、次のオプションを使用します。

1. [元に戻す] および [やり直し] オプション
タスク内でデータまたはメタデータに加えた最新の更新を元に戻します。元に戻した更新をやり直します。
タスク内の全ページに対して実行する検索および置換操作は取り消しできません。
2. タスクアクションメニュー
現在のタスクに対してアクションを実行します。ワークフロー内の次のステップにタスクデータをエクスポートするか、タスクを送信します。
3. 検索および置換オプション
入力した文字列と一致するデータ値を検索し置換します。
4. 強調表示
ステータスインジケータを使用しないレコードを表示します。
5. 検索および置換
例外レコードのデータ値を検索および置換するにはオプションを表示または非表示にします。

6. メタデータフィルタオプション
データ品質の問題のタイプやレコードステータスなどのレコードメタデータに基づいて、タスク内のレコードを表示または非表示にします。
7. レコードアクションメニュー
選択したレコードステータスを更新します。例えば、レコードの承認、拒否、またはレコードステータスのクリアです。
8. データフィルタのリセット
カラムデータフィルタとして入力した値をクリアします。
9. 注カラム
タスク所有者がレコードに追加した注を確認します。
10. レコードステータスカラム
タスク所有者がレコードに対して実行した直近のアクションのステータスインジケータを表示します。
11. レコード確認ステータスカラム
タスク所有者がレコードに定義した直近のステータスのステータスインジケータを表示します。
12. カラムデータフィルタオプション
1つ以上のデータカラム内のデータに基づいて、タスク内のレコードを表示または非表示にします。
13. レコードセレクト
レコードアクションを適用するレコードを選択します。
14. ナビゲーションオプション
タスク内の他のページに移動します。

例外タスクのエラーの修正

例外修正タスクおよび例外確認タスクのレコードを更新できます。タスク内のレコードを調べて、データ品質エラーを含むフィールドを更新します。

【**データ編集**】タブでレコードを編集します。

1. 例外タスクを開きます。
タスクに【**データ編集**】タブが表示されるのを確認します。
2. 【**編集**】をクリックします。
3. レコードを調べます。修正可能なエラーを含むデータフィールドを選択します。
注: フィールド上でポインタを動かしてフィールドが含むデータエラーのタイプを確認します。
4. データフィールドに正しい値を入力します。
フィールドの表示が変わります。Analyst ツールによりフィールドの背景の赤がクリアされ、右上隅に緑色のチェックマークが追加されます。
5. タスクを保存します。
タスク内の他のレコードにも同じ手順を実行します。

例外タスクの検証エラー

データを格納するカラムに対して無効なデータ値を入力しようとすると、Analyst ツールに検証エラーが表示されます。数値型のカラムには、文字列型のデータ値を入力できません。データ値がカラムの最大文字数を超える場合は、その値をカラムに入力することはできません。

また、Analyst ツールでは、検索および置換操作の間に検証エラーを表示することもできます。置き換える値がターゲットカラムで無効な場合は、値の置換は行えません。

例: インドの都市名

国際的組織のデータスチュワードであるあなたは、インドの都市名のうち古い都市名が使われているものを識別する例外タスクに取り組んでいます。あなたは現在の都市名と古い都市名のリストを使用して、タスク内のレコードを更新します。各都市に対して、[すべて置換] オプションを使用して名前を更新します。デフォルトのオプションを使用して、すべての文字列データ内の値を置き換えます。

都市名が格納されたカラムの 1 つで、現在の名前が収まらない長さの最大文字数が指定されています。例えば、インドのケララ州の都市 *Trivandrum* は、*Thiruvananthapuram* という名前に置き換えられます。Analyst ツールがこのカラムに含まれる名前を更新しようとすると、操作は失敗し、Analyst ツールにエラーが表示されます。

例外レコードのステータスインジケータ

レコードに対する作業が完了したら、レコードのステータスを設定します。タスク完了時、別のユーザーまたは別のデータ処理がレコードに対して実行可能なアクションがステータスに表示されます。レコードに対する作業が完了したら、レコードデータの状態を表すステータスを設定します。

[データ編集] タブの [レコードアクション] メニューからステータスインジケータを選択します。ステータスインジケータはタスク内のデータ行の横にアイコンとして表示されます。前のタスクでユーザーがレコードステータスを設定した場合、このステータスを更新またはクリアできます。

レコードに対する作業が完了したら、次のステータスインジケータのいずれかを選択します。

レコードの承認

組織データを含むストレージで承認可能なレコードであることを示します。

例外修正タスクおよび例外確認タスクでレコードを承認できます。

レコードの拒否

組織データを含むストレージで承認不可のレコードであることを示します。

データベース内のビジネスエンティティを正しく識別できないと判定されたレコードを拒否します。データベーステーブルから拒否されたレコードはダウンストリームプロセスにより削除されることがあります。

例外修正タスクおよび例外確認タスクでレコードを拒否できます。

レコードの再処理

レコードに検証できないデータ品質問題が含まれることを示します。

レコードデータの正確性を決定できない場合、レコードを再処理します。ダウンストリームタスクの他のユーザーがレコードデータの検証または更新をすることがあります。または、ダウンストリームプロセスでマッピングの分析および修正用テーブルにレコードを書き込むことがあります。

例外修正タスクおよび例外確認タスクでレコードを再処理できます。

確認済みとしてマーク

レコードデータが確認済みであることを示します。

レコードを確認済みとマークして他のユーザーにレコードの調査を行ったことを示します。このステータスは、レコードのデータ品質を説明するものではなく、レコードデータに対してさらに別のアクションを指定するものでもありません。

レコードに確認済みとマークした場合、他のステータスインジケータを設定できます。例えば、確認済みとマークしたレコードを再処理できます。

任意のタスク内でレコードを確認済みにマークできます。

レコード編集の承認

例外確認タスクでレコードを分析し、組織データを含むストレージで承認可能なレコードだと決定したことを示します。

重複確認タスクでもレコード編集の承認を行うことができます。

レコード編集の拒否

例外確認タスクでレコードを分析し、組織データを含むストレージで承認不可のレコードだと決定したことを示します。

重複確認タスクでもレコード編集の拒否を行うことができます。

レコードステータスのクリア

レコードから次のステータスインジケータをクリアします。

- レコードの承認
- レコードの拒否
- レコードの再処理

レコードの現在のステータスが正しくないと決定した場合、ステータスをクリアします。

例外修正タスクおよび例外確認タスクでステータスインジケータをクリアできます。

確認者のステータスをクリア

レコードから次のステータスインジケータをクリアします。

- レコード編集の承認
- レコード編集の拒否
- 確認済みとしてマーク

現在の確認ステータスが正しくないと決定した場合、ステータスをクリアします。例えば、前のタスクで他のユーザーが確認したレコードのステータスをクリアすることがあります。

ステータスインジケータはどのタスクでもクリアできます。

注: ステータスインジケータはレコードの現在の状態を表している必要があります。例えば、レコード内でエラーを特定したためにレコードを拒否することがあります。エラーを修正したら、レコードステータスを更新します。

例外レコードのステータスの設定

レコードに対する作業を完了したら、レコードのステータスインジケータを設定します。一度に複数のレコードに対してステータスインジケータを適用できます。

例外修正タスクおよび例外確認タスクでレコードステータスを設定できます。現在のレコードステータスに同意しない場合、ステータスをクリアできます。

1. 例外タスクを開きます。

タスクに【**データ編集**】タブが表示されるのを確認します。

2. **【編集】** をクリックします。

3. 1つ以上のレコードを選択します。

レコードを選択するには、レコードのデータ行の先頭にあるチェックボックスを選択します。タスク内の全レコードを選択するには、カラム名を示す行のチェックボックスを選択します。

すべてのレコードを選択する場合次の要素を考慮します。

- すべてのレコードを選択する操作は現在のワークスペースのレコードに適用されます。
- フィルタをタスクに適用すると、すべてのレコードを選択する操作はフィルタ条件を満たすワークスペースのレコードに適用されます。

4. **【レコードアクション】** メニューを開き、レコードに適用するステータスインジケータを選択します。

例外修正タスクおよび例外確認タスクで次のステータスインジケータを選択できます。

- レコードに有効なデータが含まれていることを示すには、**【レコードの承認】** を選択します。
- レコードに有効なデータが含まれていないことを示すには、**【レコードの拒否】** を選択します。
- レコードをさらに分析する必要があることを示すには、**【レコードの再処理】** を選択します。
- レコードを検査済みであることを示すには、**【確認済みとしてマーク】** を選択します。インジケータを別のインジケータと同時に選択することができます。

例外確認タスクで次のステータスインジケータを選択できます。

- ビジネスデータへの追加に適したレコードであることを確認するには、**【承認としてマーク】** を選択します。
- ビジネスデータへの追加に適さないレコードであることを確認するには、**【拒否としてマーク】** を選択します。

5. レコードの現在のステータスに同意しない場合、次のオプションのいずれか、または両方を選択します。

- レコードを承認、拒否、または再処理するインジケータをクリアするには、**【レコードステータスのクリア】** を選択します。
- 確認のタイプを指定するインジケータをクリアするには、**【確認済みステータスのクリア】** を選択します。

タスク内の他のレコードにも同じ手順を実行します。

注: **【レコードアクション】** メニューにはレコードにメモを追加するオプションが含まれます。加えた変更や選択したステータスインジケータについて説明するには、レコードにメモを追加します。

例外タスクのフィルタオプション

あるタスクに対して Analyst ツールが表示するレコードを、フィルタを使用して表示または非表示にします。フィルタは **【データ編集】** タブと **【データ監査】** タブで定義できます。

レコードデータ値またはレコードメタデータを使用して、フィルタを定義できます。レコードデータフィルタを使用して、1つ以上のデータ値を共有するレコードを検索します。メタデータフィルタを使用して、共通するデータ品質の問題やステータスを共有するレコードを検索します。

例えば、指定した期間内でユーザーが有効と承認した全レコードを取得するフィルタを定義する場合があります。フィルタ条件に一致するレコードが取得されます。各レコードを確認し、ユーザーが行った決定を検証できます。

各タブで複数のフィルタを適用できます。複数のフィルタを適用した場合、すべてのフィルタ条件に一致するレコードが返されます。

レコードデータフィルタ

レコードデータフィルタを使用して、指定したカラムのデータ値を共有するレコードを検索します。

各カラムにフィルタオプションがあります。フィルタを適用するには、カラムの最上部でフィルタに値を入力します。Analyst ツールに、選択したフィールドに入力した値を含むレコードが表示されます。

タスクデータの複数のカラムにフィルタを定義できます。複数のカラムにフィルタを定義した場合、Analyst ツールには、定義したすべてのフィルタに一致するレコードが表示されます。

[データ編集] タブのメタデータフィルタ

[データ編集] タブのメタデータフィルタを使用して、共通のメタデータ特性を持つレコードのセットを識別します。

[データ編集] タブには次のフィルタオプションがあります。

問題のタイプ

ワークフローがレコードデータで識別したデータ品質問題のタイプを示します。問題を選択すると、Analyst ツールにその問題を含むすべてのレコードが表示されます。

フィールドの上にポインタを移動すると、フィールドが表示するデータ品質問題が表示されます。データ品質問題を含むフィールドは [例外] ワークスペースに赤で表示されます。

優先度

マッピングタスクがレコード内のデータ品質問題に割り当てた優先度を示します。優先度は数値で表されます。優先度の値を選択すると、Analyst ツールにその優先度を含むすべてのレコードが表示されます。

ステータス

タスク内のレコードに割り当てたステータスを示します。ステータスを選択すると、Analyst ツールにそのステータスを使用するすべてのレコードが表示されます。

次のステータスオプションを選択することができます。

- 任意。ステータスに関係なく、タスク内の任意のレコード。
- 承認済み。組織データを含む永続ストレージに適したレコード。
- 拒否済み。組織データを含む永続ストレージに適さないレコード。
- 再処理済み。他のアプリケーションでさらに分析する必要のあるレコード。
- 空。現在のステータスがないレコード。

確認

現在のタスク内でのレコードの確認ステータスを示します。

確認オプションは、次のいずれかを選択できます。

- すべてのレコード。
- 承認済み。確認タスクで承認済みとユーザーがマークしたレコード。
- 拒否済み。確認タスクで拒否済みとユーザーがマークしたレコード。
- 確認済み。ユーザーが確認済みとマークしたレコード。
- 空。現在の確認ステータスがないレコード。

[データ監査] タブのメタデータフィルタ

[データ監査] タブのメタデータフィルタを使用して、指定した期間でユーザーが作業をしたレコードを識別します。共通のステータスインジケータを持つレコードも検索できます。

[データ監査] タブでは、レコードデータを編集できません。レコードデータを編集するには、[データ編集] タブを使用します。

[データ監査] タブには次のフィルタオプションがあります。

更新日

タスクデータに適用する時間を定義します。期間を定義すると、Analyst ツールはその期間内にユーザーが更新したすべてのレコードを返します。

更新者

Analyst ツールのユーザーを識別します。ユーザーを指定すると、Analyst ツールはそのユーザーが更新したすべてのレコードを返します。

ステータス

タスク内のレコードに割り当てたステータスを示します。ステータスを選択すると、Analyst ツールにそのステータスを使用するすべてのレコードが表示されます。

次のステータスオプションを選択することができます。

- 任意。ステータスに関係なく、タスク内の任意のレコード。
- 承認済み。組織データを含む永続ストレージに適したレコード。
- 拒否済み。組織データを含む永続ストレージに適さないレコード。
- 再処理済み。他のアプリケーションでさらに分析する必要のあるレコード。
- 空。現在のステータスがないレコード。
- クリア済み。ユーザーがステータスをクリアしたレコード。

確認

現在のタスク内でのレコードの確認ステータスを示します。

確認オプションは、次のいずれかを選択できます。

- すべてのレコード。
- 承認済み。確認タスクで承認済みとユーザーがマークしたレコード。
- 拒否済み。確認タスクで拒否済みとユーザーがマークしたレコード。
- 確認済み。ユーザーが確認済みとマークしたレコード。
- 空。現在の確認ステータスがないレコード。
- クリア済み。ユーザーが確認ステータスをクリアしたレコード。

例外タスクのレコードのフィルタリング

メタデータフィルタを使用して、指定したデータ特性を共有するレコードのセットを取得します。[データ編集] タブのフィルタ条件を満たすレコードを更新できます。

1. 例外修正タスクまたは例外確認タスクを開きます。

[データ編集] タブまたは [データ監査] タブを選択します。

2. [フィルタ] をクリックします。

[フィルタ] パネルが開きます。

3. タスクのデータに適用するフィルタ条件を選択します。

[データ編集] タブと [データ監査] タブには、異なるフィルタオプションセットが表示されます。

4. **[フィルタの適用]** をクリックします。

フィルタ条件に一致するレコードが取得されます。定義したフィルタ条件はタスクのクラスタリストの上に表示されます。

注: フィルタを適用後 [データ編集] タブの全レコード選択した場合、現在のワークスペースビューのフィルタ条件を満たすすべてのレコードを選択しています。タスク内のすべてのレコードを選択しません。

例外レコードでの値の検索および置換

タスク内のレコードを、指定したデータ値で検索できます。その値を他の値で置換できます。1つのレコードの値を検索して置換することも、一度に複数のレコードの値を検索して置換することもできます。

検索と置換オプションを設定するときには、名前によって1つのカラムを選択できます。または、日付、文字列、または数値データ型を使用するすべてのカラムを検索するオプションを設定できます。

データ型によってタスクを検索する場合、そのデータ型を使用するすべてのカラムで、入力した値が検索されます。文字列データカラムに含まれる値は、数値や日付も含めてすべて検索および置換できます。文字列カラムデータでは、大文字と小文字を区別して検索を実行できます。検索値と文字列データカラムのフィールドの内容の間で部分一致または完全一致を検索できます。部分一致が見つかったら、レコードフィールド内の、入力した文字に一致するすべてのインスタンスが置換されます。

以下の表に、名前データの文字列に対する部分一致と置換操作の結果を示します。

検索値	置換値	検索する文字列	結果
SMITH	SMYTH	JOHN SMITH AND MARY SMITH	JOHN SMYTH AND MARY SMYTH

日付を検索する場合は、日付カラムで使用されている日付およびタイムスタンプの形式で日付値を入力します。日付カラムの NULL 値を指定した日付値に置換できます。日付データカラムまたは数値データカラムの日付値の一部や数値の一部を検索することはできません。日付データカラムまたは数値データカラムの内容に対して大文字と小文字を区別して検索を実行することはできません。

タスクデータをフィルタリングした場合の値の検索および置換が可能です。検索および置換操作はフィルタ条件を満たすレコードに適用されます。フィルタリングされたデータを検索すると、フィルタ条件を満たすレコードが含まれるタスクデータの最初のページが表示されます。

[検索&置換] をクリックして、タスク内でオプションを表示または非表示にします。

例外レコードでの単一値の置換

指定したカラムまたは任意の日付、文字列、または数値カラム内の単一のデータ値を検索し置換します。フィルタを使用して、タスクに含まれるレコードのサブセット内の値を検索して置換します。

1. 例外タスクを開きます。
2. **[データ編集]** タブで **[編集]** をクリックします。
3. **[検索&置換]** をクリックします。

Analyst ツールでレコードのリストの上に検索と置換のオプションが表示されます。

4. **【検索】** フィールドにデータの値を入力します。
5. 検索するカラムを指定します。または、データ型を指定します。
データ型を指定すると、そのデータ型を使用するカラムが検索されます。デフォルトでは、すべての文字列データカラム内のデータ値が、Analyst ツールによって検索されます。
6. 文字列データカラム内のデータ値を検索する場合、次のオプションを選択またはクリアします。
 - 大文字小文字の区別:文字列データに対して大文字と小文字を区別して検索を実行します。
 - フィールド全体に一致。入力した値とレコードのフィールドの内容全体との一致を検索します。日付データまたは数値データのフィールドでは大文字小文字を区別する検索または部分データ検索を実行できません。
7. **【検索】** をクリックします。
指定した値が含まれる最初のレコードが強調表示されます。矢印を使用してレコード間を移動します。
8. 必要に応じて、**【置換後の文字列】** フィールドに値を入力します。
9. 強調表示された値を入力した値で置換するには、**【置換】** をクリックします。
タスク内の値の追加のインスタンスを検索および置換する場合にも同じ手順を実行します。

例外レコードでの複数の値の置換

〔すべて置換〕オプションを使用すると、1つの値の複数インスタンスが1回の操作で置換されます。現在のページまたはすべてのページのレコードや、選択したレコードに含まれる値を検索して置換できます。フィルタを使用して、タスクに含まれるレコードのサブセット内の値を検索して置換します。

1. 例外タスクを開きます。
2. 必要に応じて、複数のレコードを選択します。
複数のレコードを選択すると、〔すべて置換〕操作が、タスク内のレコードのセットに制限されます。例えば、定義したフィルタ条件と一致するすべてのレコードを選択できます。
3. **【データ編集】** タブで **【編集】** をクリックします。
4. **【検索&置換】** をクリックします。
Analyst ツールでレコードのリストの上に検索と置換のオプションが表示されます。
5. **【検索】** フィールドにデータの値を入力します。
6. 検索するカラムを指定します。または、データ型を指定します。
7. **【置換後の文字列】** フィールドに値を入力します。
8. **【すべて置換】** をクリックします。
次のいずれかのオプションを選択します。
 - **〔このページで置換〕**。現在のページのレコードに含まれる、入力した値の全インスタンスが置き換えられます。
 - **〔このページで置換〕**。タスクのすべてのレコードに含まれる、入力した値の全インスタンスが置き換えられます。
 - **〔選択したすべてで置換〕**。選択したレコードに含まれる、入力した値の全インスタンスが置き換えられます。

レコードへのメモの追加

レコードに対して実行するアクションについての説明となるメモを追加します。このメモは「データ編集」タブと「データ監査」タブに表示できます。

1. レコードを含むタスクを開きます。
2. 「例外」ワークスペースで「データ編集」タブを開きます。
3. 「編集」をクリックします。
4. チェックボックスオプションを使用してレコードを選択します。
複数のレコードを選択できます。複数のレコードを選択すると、選択したすべてのレコードに同じメモを追加することになります。
5. 「レコードアクション」メニューを開いて、「注の追加」を選択します。
Analyst ツールにテキスト入力ダイアログボックスが開きます。
6. メモのテキストをダイアログボックスに入力します。

確認中の例外タスク

例外確認タスクでレコードを確認することは、前のタスクで他のユーザーが行った作業を検証することです。例外修正タスクおよび例外確認タスクで同じ操作を実行できます。

タスク内のレコードを調べて、前のユーザーが各レコードに割り当てたステータスを確認します。レコードにエラーを見つけたら、レコードを更新します。レコードの現在のステータスに同意しない場合、ステータスを更新またはクリアします。

「レコードアクション」メニューオプションを使用して、レコードステータスを更新します。タスクに対する作業が完了したら、「タスクアクション」メニューオプションを使用してタスクステータスを更新します。

例外確認タスクでの作業

タスク内のレコードが可能な限り正確で完全であることを確認します。

1. 例外確認タスクを開きます。
2. レコードデータを検証し、各レコードに前のユーザーが割り当てたステータスを検証します。
3. レコードの更新、またはレコードステータスの変更を決定したら、「編集」をクリックします。
エラーを修正するには、エラーを含むフィールドをクリックして正しいデータ値を入力します。
レコードステータスを更新するには、次のいずれかのアクションを実行します。
 - レコードに有効なデータが含まれていることを示すには、「レコードの承認」を選択します。
 - レコードに有効なデータが含まれていないことを示すには、「レコードの拒否」を選択します。
 - レコードをさらに分析する必要があることを示すには、「レコードの再処理」を選択します。
 - レコードを検査済みであることを示すには、「確認済みとしてマーク」を選択します。インジケータを別のインジケータと同時に選択することができます。
 - ビジネスデータへの追加に適したレコードであることを確認するには、「承認としてマーク」を選択します。
 - ビジネスデータへの追加に適さないレコードであることを確認するには、「拒否としてマーク」を選択します。

- レコードを承認、拒否、または再処理するインジケータをクリアするには、[レコードステータスのクリア] を選択します。
 - 確認のタイプを指定するインジケータをクリアするには、[確認済みステータスのクリア] を選択します。
4. タスクを保存します。
- タスクに対する作業が完了したら、タスクのステータスを更新します。タスクのステータスは、レコードがワークフローの次のステージに進める状態になったことを示します。

例外タスクのルールおよびガイドライン

データセット内における例外レコードの確認と更新の処理は共同作業です。他のユーザーが前のタスクで分析したレコードに対して作業することがあります。または、作業したレコードが、タスク完了後、他のユーザーへ渡されることがあります。各ユーザーは、他のユーザーの作業を確認および更新できます。

例外レコードタスクで作業する場合、次のルールおよびガイドラインを考慮してください。

- レコードのステータスを設定すると、レコードが確認済みであることを示すことができます。この確認ステータスは、レコードの正確性やデータ品質を説明するものではありません。例えば、一連のレコードで承認、拒否、および再処理を実行でき、各レコードを確認済みとしてマークできます。
ベストプラクティスとして、調査したすべてのレコードを確認済みとしてマークします。このステータスは、ダウンストリームタスクのユーザーに、他のユーザーがレコードを調査したことを示します。レコード内のデータを更新するときは、レコードに対する別のステータスインジケータの有無に関係なく、レコードを確認済みとしてマークします。
- レコードを更新する場合、任意のステータスインジケータを設定できます。例えば、再処理するレコードを更新できます。更新を行うと、次のユーザーまたはダウンストリームデータ処理がデータを分析および修正するのに役立ちます。
- レコードのデータ値およびステータスインジケータは、複数のタスクで互いに独立して変更可能です。レコードを更新する場合、データ値とステータスインジケータが最新で正確であることを確認します。変更を行うと、他のユーザーが行った決定を無効にすることができます。
- 監査証跡にユーザーがレコードのデータ値およびステータスインジケータに加えた変更がすべて保存されます。監査証跡にはレコードに追加されたメモのテキストの変更は保存されません。
- 作業したデータは、データの修正タスクやデータの確認タスクに渡すことができます。例えば、ワークフロー内のヒューマンタスクを設定する開発者は、複数の例外修正タスクを順番に指定することができます。開発者は重複修正タスクと共に重複確認タスクを実行することができます。

第 3 章

重複レコードタスク

この章では、以下の項目について説明します。

- [重複レコードタスクの概要, 31 ページ](#)
- [優先レコード, 32 ページ](#)
- [重複レコードタスクのタイプ, 32 ページ](#)
- [\[例外\] ワークスペースの重複レコードタスク, 33 ページ](#)
- [クラスタの編集, 35 ページ](#)
- [重複レコードのステータスインジケータ, 36 ページ](#)
- [複数のクラスタでのレコードの検索, 37 ページ](#)
- [クラスタ間でのレコードの移動, 38 ページ](#)
- [空のクラスタへのレコードの移動, 39 ページ](#)
- [クラスタへのメモの追加, 39 ページ](#)
- [重複タスクのフィルタオプション, 40 ページ](#)
- [重複確認タスク, 41 ページ](#)
- [重複レコードタスクのルールおよびガイドライン, 43 ページ](#)

重複レコードタスクの概要

重複レコードタスクでは、重複または冗長情報が含まれている可能性がある、データセット内のレコードを特定します。このタスクでは、レコードが一連のクラスタとして表示されます。各クラスタは、類似するまたは同一のデータ値が含まれるレコードのセットを識別します。

重複レコードタスクの作業を行うときに、各クラスタを確認して、クラスタ内のレコードが互いに重複していないかどうかを判別します。レコードが重複している場合、クラスタで表されるレコードの優先バージョンを定義します。クラスタに一意のレコードが含まれている場合、そのレコードを空のクラスタに移動します。レコードが、別のクラスタのレコードと一致する場合は、レコードを一方のクラスタに移動できます。

すべてのクラスタを確認し、各クラスタの優先レコードを定義したら、タスクを完了します。

注: 2 つ以上のレコードが重複するのは、それらがソースデータセット内の同じエンティティを表している場合です。複数のレコードに類似データが入っていても、それらのレコードが業務にとって同じ情報を表してはいない可能性があります。組織で、重複レコードを識別する業務ルールを定義できます。

優先レコード

クラスタに対する作業を行うときに、クラスタで表される最も正確かつ完全なバージョンのレコードを作成または確認します。作成または確認したレコードが優先レコードになります。

クラスタ内の最初の行に優先レコードのデータが含まれます。優先レコードを更新するには、クラスタ内の他のレコードから優先レコードにデータ値を昇格します。

デフォルトでは、Analyst ツールにより、クラスタ内の最初のレコードのデータが優先データ行に入力されます。優先データ行が強調表示されます。フォルトデータが含まれるレコードも強調表示されます。データ値を優先レコードに昇格すると、昇格した値は強調表示されます。優先レコードに追加した値は編集できます。クラスタ内の別のレコードの値を編集することはできません。デフォルトのレコードが正しい場合、そのレコードを受け入れることができます。

注: 優先レコードはクラスタのメンバではありません。優先レコードは、ワークフローによって重複レコードデータベース内に作成される一意のレコードです。

重複レコードタスクのタイプ

ワークフローで重複修正タスクまたは重複確認タスクをユーザーに割り当てることができます。

重複修正タスクの作業を行うときは、タスク内の各クラスタのレコードを調べます。クラスタを開き、クラスタ内の優先レコードに、利用可能な最も正確なデータが含まれていることを確認します。重複確認タスクの作業を行うときは、前のタスクでクラスタごとに他のユーザーが行った決定が適切であることを確認します。

各タスクタイプのワークスペースレイアウトは同じです。各タスクタイプで使用するワークスペースオプションは似ています。クラスタデータおよびステータスインジケータの更新は、どちらのタスクタイプでも同じ方法を使用して行います。重複確認タスクの作業を行うときは、追加のオプションを使用して前のユーザーの作業を承認または拒否できます。

重複の修正手順

〔開始〕 ワークスペースから重複確認タスクを開きます。〔例外〕 ワークスペースの〔データ編集〕タブでタスクが開きます。

重複修正タスクの作業を行うときは、次の手順を実行します。

1. タスクを開き、編集モードに入ります。
Analyst ツールには、タスク内の最初のクラスタが表示されます。現在のクラスタの作業を行ったり、別のクラスタを開いたりできます。フィルタオプションを使用して、指定したデータ値またはステータスを含むクラスタを表示または非表示にします。
2. 優先レコードと現在のクラスタ内の他のレコードを調べます。
別のレコードのフィールドに優先レコードの同じフィールドよりも正確な値が含まれている場合、その値を優先レコードに昇格します。
3. 必要に応じて、以下のいずれかのタスクを実行し、クラスタに最も正確な値が含まれていることを確認します。
 - 優先レコードの値を編集する。
 - 他のクラスタで、似た値が含まれるレコードを探す。
 - クラスタ間でレコードを移動する。

- クラスタを作成し、そのクラスタにレコードを移動する。
4. クラスタステータスを更新し、クラスタを確認したことを示します。
 5. タスクのすべてのクラスタに対する作業が完了したら、タスクステータスを更新します。

重複の確認手順

〔開始〕 ワークスペースから重複修正タスクを開きます。〔例外〕 ワークスペースの〔データ編集〕タブでタスクが開きます。

例外確認タスクの作業を行うときは、次の手順を実行します。

1. タスクを開き、編集モードに入ります。
Analyst ツールには、タスク内の最初のクラスタが表示されます。現在のクラスタの作業を行ったり、別のクラスタを開いたりできます。フィルタオプションを使用して、指定したデータ値またはステータスを含むクラスタを表示または非表示にします。
2. 各クラスタのコンテンツを確認します。別のタスクでユーザーがクラスタに割り当てたステータスを確認または更新します。
フィルタを使用して、ユーザーが確認したまたは確認していないクラスタを探します。
 - クラスタのコンテンツとステータスに同意する場合、変更は加えません。
 - いずれかのクラスタのコンテンツに同意しない場合、クラスタを更新します。
 - いずれかのクラスタのステータスに同意しない場合、ステータスを更新します。
3. タスクのすべてのクラスタに対する作業が完了したら、タスクステータスを更新します。

〔例外〕 ワークスペースの重複レコードタスク

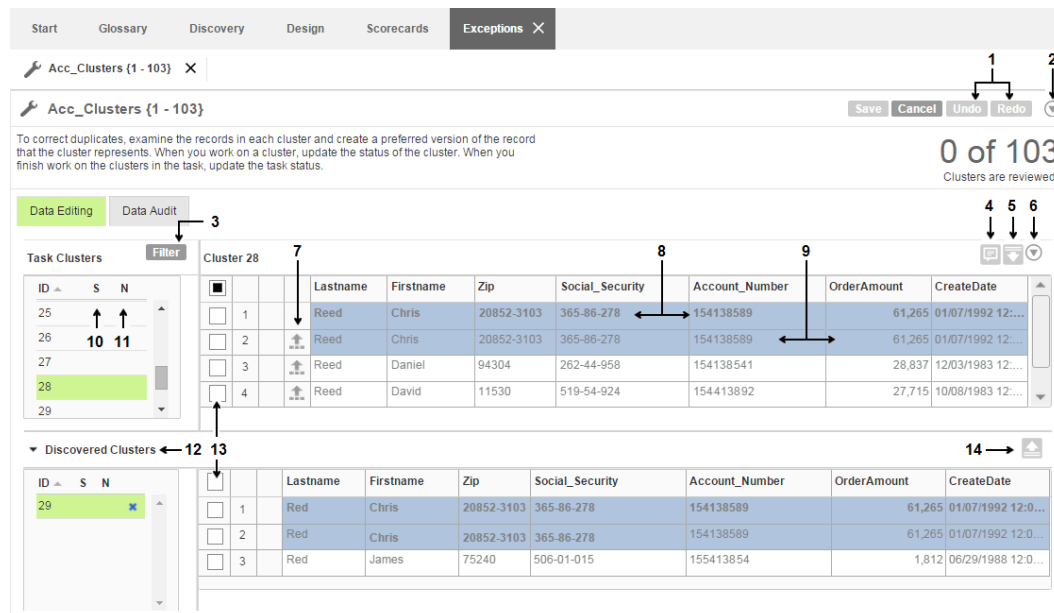
重複レコードタスクを開くと、〔例外〕ワークスペースに、タスク内のクラスタのリストが表示され、タスク内の最初のクラスタのレコードが表示されます。タスクは〔データ編集〕タブに表示されます。タブには、重複修正タスクおよび重複確認タスク用の共通のオプションセットが表示されます。

ワークスペースの左側にタスク内のクラスタのリストが表示されます。クラスタを選択し、その中に含まれているレコードを表示します。クラスタ内の最初のレコードが現在の優先レコードです。クラスタ内の他のレコードを確認し、そのレコードに、優先レコードに属しているデータが含まれているかどうかを判断します。レコードの値を優先レコードに追加するには、値をクリックします。優先レコードを別のレコードで置き換えるには、レコードを選択し、レコードを昇格するためのカラムオプションを使用します。

値を優先レコードに追加すると、ワークスペースは追加された値を強調表示します。優先レコードを別のレコードで置き換えると、ワークスペースは昇格された値を強調表示します。タスクに加えたすべての変更のリストを表示するには、〔データ監査〕ビューを選択します。

クラスタに対する作業が完了したら、クラスタのステータスを更新します。タスク内のクラスタに対する作業が完了したら、タスクステータスを更新し、ワークフローの次のステージに向けてクラスタの準備ができたことを示します。

次の画像は、クラスタが含まれるタスクを編集モードで開いたときの〔データ編集〕タブを示しています。



タスクに対する作業を行うときは、次のオプションを使用します。

1. [元に戻す] および [やり直し] オプション
タスク内でデータまたはメタデータに加えた最新の更新を元に戻します。元に戻した更新をやり直します。
2. タスクアクションメニュー
現在のタスクに対してアクションを実行します。ワークフロー内の次のステップにタスクデータをエクスポートするか、タスクを送信します。
3. フィルタのオプション
指定した条件を満たすクラスタをタスク内で探します。フィルタを適用する場合、現在のクラスタを閉じて、フィルタ条件を満たす最初のクラスタを開きます。
フィルタは、検出されたクラスタにも適用されます。フィルタを適用すると、検出されたリスト内でフィルタ条件を満たすクラスタが [検出されたクラスタ] リストに表示されます。
4. メモのオプション
[タスククラスタ] リストの現在のクラスタにユーザーが追加したメモを確認します。
5. レコードを下クラスタに移動する
選択した1つ以上のレコードを上クラスタから下クラスタに移動します。[検出されたクラスタ] リストからクラスタを開く場合、このオプションを使用します。
6. クラスタアクションメニュー
[タスククラスタ] リストの現在のクラスタのステータスを更新します。クラスタにメモを追加したり、クラスタを確認済みとしてマークしたり、クラスタの確認ステータスをクリアしたりできます。
7. 昇格オプション
優先レコードのデータを、選択した行のデータで更新します。
8. 優先レコードデータ
クラスタの優先レコードの現在のデータ。
9. 優先レコードデータのソースレコード。
優先レコードのソースデータを提供した、クラスタ内のレコード。デフォルトでは、優先レコードはシーケンス ID 値を持つ、クラスタ内のレコードです。
10. クラスタの確認ステータスカラム
クラスタの現在のステータスに対するステータスインジケータを表示します。

11. クラスタのメモカラム
どのクラスタにユーザーがメモを追加したかを示します。
12. [検出されたクラスタ] オプション
指定した 1 つ以上のデータ値が含まれるクラスタを探したり、クラスタを作成したりします。
13. レコード選択オプション
1 つ以上のレコードを選択します。
14. レコードを上クラスタに移動する
下のクラスタから選択した 1 つ以上のレコードを上クラスタに移動します。 [検出されたクラスタ] リストからクラスタを開く場合、このオプションを使用します。

クラスタの編集

優先レコードとクラスタ内の他のレコードを調べます。クラスタで表わされるビジネスエンティティにとって最も正確なデータが優先レコードに含まれていることを確認します。

現在の優先レコードが正しい場合、変更は加えません。

1. 重複修正タスクを開きます。
タスクのクラスタのリストが [データ編集] タブのカラムに表示されます。デフォルトでは、最初のクラスタが開きます。
2. 必要に応じて、別のクラスタを選択します。
例えば、現在のクラスタのステータスが [確認済み] の場合、別のクラスタを選択します。
3. クラスタ内のレコードを調べます。クラスタで表わされるビジネスエンティティにとって最も正確なデータが優先レコードに含まれているかどうかを判断します。
優先レコードはクラスタ内の最初のレコードです。Analyst ツールにより、優先レコードと優先データのソースレコードが強調表示されます。
4. [編集] をクリックします。
5. 優先レコードを更新するには、次のいずれかのアクションを実行します。
 - 別のレコードから優先レコードにデータ値を昇格する。
値を昇格するには、ソースレコードのデータフィールドをクリックします。選択した値によって、優先レコードの同じカラムの値が置き換えられます。
 - 優先レコードのデータ値を更新する。
正確なバージョンのデータが含まれているレコードがクラスタ内にない場合、優先レコードを更新します。
 - 現在の優先レコードをクラスタ内の別のレコードで置き換える。
優先レコードを置き換えるには、そのレコードが含まれる行で昇格ツールをクリックします。
6. クラスタをタスクに保存します。
クラスタでの作業が完了したら、クラスタのステータスを [確認済み] に更新します。

関連項目：

- [「複数のクラスタでのレコードの検索」 \(ページ 37\)](#)
- [「空のクラスタへのレコードの移動」 \(ページ 39\)](#)

重複レコードのステータスインジケータ

クラスタに対する作業が完了したら、クラスタのステータスを設定します。このステータスは、タスクを完了したときに別のユーザーまたは別のデータ処理がクラスタデータに対して実行できるアクションを示します。クラスタに対する作業が完了したら、優先レコードの状態を表すステータスを設定します。

[データ編集] タブの [クラスタアクション] メニューからステータスインジケータを選択します。ステータスインジケータがクラスタのリスト内にアイコンとして表示されます。前のタスクでユーザーがクラスタステータスを設定していた場合は、ステータスを更新またはクリアできます。

クラスタに対する作業が完了したら、次のステータスインジケータのいずれかを選択します。

確認済みとしてマーク

クラスタデータを確認したことを示します。

クラスタを確認済みとしてマークすると、クラスタを調べたことが他のユーザーに示されます。このステータスは優先レコードのステータスを説明するものではありません。また、レコードデータに対するさらなるアクションを指定するものでもありません。

クラスタを確認済みにマークしても、レコードに別のステータスインジケータを設定することができません。例えば、確認済みとしてマークした優先レコードを拒否できます。

クラスタはどのタスクでも確認済みとしてマークできます。

確認者のステータスをクリア

クラスタからすべてのステータスインジケータをクリアします。現在のステータスが正しくないと判断した場合、ステータスをクリアします。

ステータスインジケータはどのタスクでもクリアできます。

レコード編集の承認

重複確認タスクでクラスタを分析したこと、および現在の優先レコードに同意することを示します。クラスタで表わされるビジネスエンティティにとって最も正確なデータがレコードに含まれていると判断した場合、レコードを承認します。

例外確認タスクでのレコード編集を承認することもできます。

レコード編集の拒否

重複確認タスクでクラスタを分析したこと、および現在の優先レコードに同意しないことを示します。クラスタで表わされるビジネスエンティティにとって最も正確なデータがレコードに含まれていないと判断した場合、レコードを拒否します。

例外確認タスクでのレコード編集を拒否することもできます。

注: ステータスインジケータは、優先レコードの現在の状態を表している必要があります。例えば、優先レコード内でエラーが確認されたため、優先レコードを拒否する場合があります。このエラーを修正する場合、クラスタのステータスを更新します。

クラスタのステータスの設定

クラスタに対する作業が完了したら、クラスタのステータスを設定します。クラスタ内の現在のレコード状態を表しているステータスを設定します。ステータスは、ワークフローの次のステージで別のユーザーまたは別のデータ処理がデータに対して実行するアクションを決定します。

クラスタのステータスを設定するには、**【クラスタアクション】** メニューの各オプションを使用します。

1. 重複修正タスクまたは重複確認タスクを開きます。
タスクのクラスタのリストが「データ編集」タブのカラムに表示されます。デフォルトでは、最初のクラスタが開きます。
2. 必要に応じて、別のクラスタを選択します。
3. **【編集】** をクリックします。
4. 必要に応じて、優先レコードを更新したり、クラスタのコンテンツを更新したりします。
5. クラスタのステータスを更新するには、次のアクションを 1 つ以上実行します。
 - クラスタを確認済みとしてマークする。
クラスタを確認済みとしてマークすると、クラスタを調べたことが他のユーザーに示されます。重複修正タスクと重複確認タスクでクラスタを確認済みとしてマークできます。
 - 確認者のステータスをクリアする。
ステータスをクリアすると、クラスタからすべてのステータスインジケータが削除されます。重複修正タスクと重複確認タスクで確認者のステータスをクリアできます。
 - クラスタ内の優先レコードを承認する。
レコードを承認すると、クラスタ内の現在の優先レコードに同意することが示されます。重複確認タスクで優先レコードを承認できます。
 - クラスタ内の優先レコードを拒否する。
レコードを承認すると、クラスタ内の現在の優先レコードに同意しないことが示されます。重複確認タスクで優先レコードを拒否できます。
6. クラスタをタスクに保存します。
タスク内のすべてのクラスタを確認したら、タスクのステータスを更新します。このタスクのステータスは、ワークフローの次のステージに向けてクラスタのレコードの準備ができたことを示します。

複数のクラスタでのレコードの検索

指定した 1 つ以上のデータ値が含まれるレコードを他のクラスタで検索できます。レコードを検索するには、**【検出されたクラスタ】** オプションを使用します。

現在のクラスタのレコードに一致する可能性があるレコードを他のクラスタで検索します。組織データセット内のビジネスエンティティと同じビジネスエンティティをレコードが表している場合、レコードを同じクラスタに移動します。

1. 重複修正タスクを開きます。
タスクのクラスタのリストが「データ編集」タブのカラムに表示されます。デフォルトでは、最初のクラスタが開きます。
2. 必要に応じて、別のクラスタを選択します。
3. **【検出されたクラスタ】** オプションを展開します。
デフォルトでは、このオプションにクラスタデータは表示されません。

4. **【クラスタの検索】** をクリックします。
5. 探すデータの値が入ったカラムを選択します。
6. 検索するデータ値を入力します。
レコードカラムに表示されるデータ値を入力するか、テキストパターンでワイルドカード文字を入力します。ワイルドカード文字には、アスタリスク (*) またはパーセント記号 (%) を使用できます。
7. 必要に応じて、他のカラムのデータ値を入力します。
8. **【検索】** をクリックします。
検索対象の値を含んだ、タスクデータ内のレコードが返されます。複数の値を検索した場合、各値を含むレコードが返されます。
9. 検索結果の 1 つ以上のレコードを選択し、**【開く】** をクリックします。
[データ編集] タブに、選択したレコードを含んだクラスタが表示されます。
10. [データ編集] タブでクラスタをクリックして開きます。
11. クラスタ内のレコードを比較します。
[データ編集] タブ内のクラスタを更新するには、**【編集】** をクリックします。

関連項目：

- [「空のクラスタへのレコードの移動」 \(ページ 39\)](#)
- [「クラスタの編集」 \(ページ 35\)](#)

クラスタ間でのレコードの移動

2 つ以上のクラスタに同じエンティティを表すレコードが含まれる場合、そのレコードをクラスタ間で移動します。他のクラスタの優先レコードにデータを提供する可能性があるレコードを移動します。

[データ編集] タブでクラスタを開き、クラスタ間でレコードを移動できます。

1. 重複修正タスクを開きます。
タスクのクラスタのリストが [データ編集] タブのカラムに表示されます。デフォルトでは、最初のクラスタが開きます。
2. 必要に応じて、別のクラスタを選択します。
例えば、現在のクラスタのステータスが [確認済み] の場合、別のクラスタを選択します。
3. **【編集】** をクリックします。
4. 現在のクラスタ内のレコードに一致するレコードを別のクラスタ内で検索します。
レコードを検索するには、**【検出されたクラスタ】** オプションを使用します。この検索操作により、複数のクラスタ内のレコードが返される場合があります。
5. 検索しているレコードが含まれるクラスタを開きます。
6. クラスタ内のレコードを比較します。
7. あるクラスタのレコードが別のクラスタのレコードとより一致する場合、レコードをそのクラスタに移動します。
レコードを移動するには、**【レコードの移動】** アイコンを使用します。
8. クラスタをタスクに保存します。

空のクラスタへのレコードの移動

タスク内のどのクラスタにも属していない有効なレコードを特定した場合、レコードを空のクラスタに移動します。作成したクラスタにレコードを移動します。

有効なレコードには、組織データセットに属しているビジネスエンティティに関する情報が含まれます。レコードは一意である場合や、別のクラスタのレコードと一致する場合があります。作成したクラスタには複数のレコードを追加できます。

1. 重複修正タスクを開きます。
タスクのクラスタのリストが「データ編集」タブのカラムに表示されます。デフォルトでは、最初のクラスタが開きます。
2. 必要に応じて、別のクラスタを選択します。
3. クラスタ内のレコードを調べます。
クラスタ内の他のレコードと一致しない有効なレコードがクラスタに含まれていることを確認します。
4. 「編集」をクリックします。
5. 「クラスタの検出」オプションを展開します。
6. 「クラスタの作成」をクリックします。
現在のクラスタの下に空のクラスタが表示されます。
7. クラスタに追加するレコードを選択します。
8. レコードをクラスタに移動します。
そのレコードが、作成したクラスタの優先レコードになります。
注: 現在のクラスタ内の「注」アイコンの横に「レコードの移動」アイコンがあります。
9. 優先レコードに一致する他のレコードを、作成したクラスタに移動します。

関連項目：

- [「複数のクラスタでのレコードの検索」 \(ページ 37\)](#)
- [「クラスタの編集」 \(ページ 35\)](#)

クラスタへのメモの追加

メモを追加し、クラスタに対して実行したアクションについて説明します。このメモは「データ編集」タブと「データ監査」タブに表示できます。

1. 「例外」ワークスペースで「データ編集」タブを開きます。
2. クラスタを開きます。
3. 「クラスタアクション」メニューで、「注の追加」を選択します。
「注」ボックスが開きます。
4. 「注」ボックスにコメントを入力します。

重複タスクのフィルタオプション

フィルタを使用して、指定したデータ条件に一致するクラスタをタスクから取得できます。フィルタは「データ編集」タブと「データ監査」タブで定義できます。

フィルタオプションを使用して、指定したデータ値を含むまたは指定したステータスを共有するクラスタを表示または非表示にします。「データ監査」タブには、クラスタの保存日と確認ステータスに基づいてクラスタを表示または非表示にする追加のオプションが含まれます。

各タブで複数のフィルタを適用できます。複数のフィルタを適用した場合、すべてのフィルタ条件に一致するレコードが返されます。

重複レコードタスクでのデータフィルタ

「データ編集」タブのフィルタを使用して、共通のデータ機能を持つクラスタのセットを特定します。フィルタを適用した場合、フィルタ条件に一致するクラスタのリストが表示されます。

「データ編集」タブには次のフィルタオプションがあります。

値

選択したカラムの値を持つ 1 つ以上のレコードを含むクラスタが返されます。カラム名を選択し、値を入力します。

ステータス

選択したステータスを表示するクラスタが返されます。

次のステータスインジケータを選択できます。

- 任意。ステータスに関係なく、タスク内のすべてのクラスタを示します。
- 承認済み。組織データを含むストレージに適した優先レコードが含まれるクラスタ。
- 拒否済み。組織データを含むストレージに適した優先レコードが含まれないクラスタ。
- 確認済み。確認したクラスタ。このステータスは、優先レコードのステータスを示すものではありません。
- 空。現在のステータスを持たないクラスタ。

重複レコードタスクでの監査証跡フィルタ

「データ監査」タブのフィルタを使用して、クラスタのメタデータに基づいてクラスタを表示または非表示にします。日付、ユーザー、ステータスでクラスタをフィルタリングできます。

「データ監査」タブには次のフィルタオプションがあります。

更新日

指定した期間にユーザーが更新したクラスタが返されます。

更新者

ユーザーが更新したクラスタが返されます。Analyst ツールのユーザーを識別します。ユーザーでフィルタリングするには、ユーザー名を入力します。

ステータス

指定したステータスを含むクラスタが返されます。

次のステータスオプションを選択することができます。

- 任意。ステータスに関係なく、タスク内のすべてのクラスタを示します。
- クラスタに移動されました。別のクラスタに移動されたレコードを含んでいたすべてのクラスタ。
- クラスタから移動されました。別のクラスタから移動されたレコードを含んでいるすべてのクラスタ。
- 空。現在のステータスを持たないクラスタ。

確認

指定した確認ステータスが含まれるクラスタが返されます。

次の確認オプションを選択することができます。

- すべてのクラスタ。確認ステータスに関係なく、タスク内のすべてのクラスタ。
- 確認済み。ユーザーが確認済みとしてマークしたクラスタ。
- 承認済み。確認タスクでユーザーが承認済みとしてマークしたクラスタ。
- 拒否済み。確認タスクでユーザーが拒否済みとしてマークしたクラスタ。
- 空。現在の確認ステータスを持たないクラスタ。
- クリア済み。ユーザーがクリアした以前の確認ステータスを持つクラスタ。

重複レコードタスクでのクラスタのフィルタリング

フィルタを使用して、指定したデータ特性を共有するレコードのセットを取得します。[データ編集] タブでフィルタを適用すると、フィルタ条件に一致する 1 つ以上のレコードが含まれるクラスタが表示されます。[データ監査] タブでフィルタを適用すると、フィルタ条件に一致するすべてのレコードが表示されます。

フィルタにより [データ編集] タブに返されるクラスタには、フィルタ条件に一致しないレコードが含まれる場合があります。フィルタを使用すると、条件を満たす少なくとも 1 つのレコードが含まれるクラスタが返されます。

1. 重複修正タスクまたは重複確認タスクを開きます。
2. [データ編集] タブまたは [データ監査] タブを選択します。
3. [フィルタ] をクリックします。

[フィルタ] パネルが開きます。

注: [データ編集] タブと [データ監査] タブには、異なるフィルタオプションセットが表示されます。

4. タスクのデータに適用するフィルタ条件を選択します。
5. [フィルタの適用] をクリックします。

フィルタ条件に一致するレコードが取得されます。

定義したフィルタ条件はタスクのクラスタリストの上に表示されます。

重複確認タスク

例外確認タスクでクラスタを確認するときは、前のタスクで他のユーザーが行った作業を検証します。重複確認タスクおよび重複修正タスクと同じ操作を実行できます。

クラスタ内のレコードを調べて、優先レコードが、クラスタ内の最も正確なバージョンのレコードであることを確認します。前のユーザーの決定に同意しない場合、優先レコードとクラスタステータスを更新できます。

レコードのステータスを更新するには、**【クラスタアクション】** メニューオプションを使用します。タスクに対する作業が完了したら、**【タスクアクション】** メニューオプションを使用してタスクステータスを更新します。

注: クラスタ内の他のレコードに、業務で維持するレコードが含まれていないことを確認します。別のクラスタにレコードを移動したり、クラスタを作成してレコードを保存したりできます。

重複確認タスクでの作業

各クラスタ内の優先レコードが、クラスタ内の最も正確なバージョンのレコードを表していることを確認します。

1. 重複確認タスクを開きます。

タスクのクラスタのリストが **【データ編集】** タブのカラムに表示されます。デフォルトでは、最初のクラスタが開きます。

2. クラスタ内のレコードを調べます。

クラスタ内のデータの最も正確なバージョンが優先レコードに入っているか確かめます。

クラスタステータスが **【確認済み】** であることを確認します。

3. クラスタのコンテンツまたはクラスタステータスを更新するには、**【編集】** をクリックします。

優先レコードを更新するには、次のアクションを 1 つ以上実行します。

- 別のレコードから優先レコードにデータ値を昇格する。

値を昇格するには、ソースレコードのデータフィールドをクリックします。選択した値によって、優先レコードの同じカラムの値が置き換えられます。

- 優先レコードのデータ値を更新する。

正確なバージョンのデータが含まれているレコードがクラスタ内にない場合、優先レコードを更新します。

- 現在の優先レコードをクラスタ内の別のレコードで置き換える。

優先レコードを置き換えるには、そのレコードが含まれる行で昇格ツールをクリックします。

レコードを別のクラスタに移動したり、レコードを別のクラスタからインポートしたりすることもできます。クラスタ間でレコードを移動するには、**【検出されたクラスタ】** オプションを使用します。

クラスタのステータスを更新するには、次のいずれかのアクションを実行します。

- クラスタを確認済みとしてマークする。

クラスタを確認済みとしてマークすると、クラスタを調べたことが他のユーザーに示されます。

- クラスタ内の優先レコードを承認する。

レコードを承認すると、クラスタ内の現在の優先レコードに同意することが示されます。

- クラスタ内の優先レコードを拒否する。

レコードを承認すると、クラスタ内の現在の優先レコードに同意しないことが示されます。

- 確認者のステータスをクリアする。

ステータスをクリアすると、クラスタからすべてのステータスインジケータが削除されます。

4. クラスタに対する作業が完了したら、クラスタをタスクに保存します。

タスク内のすべてのクラスタを確認したら、タスクのステータスを更新します。このタスクのステータスは、ワークフローの次のステージに向けてクラスタのレコードの準備ができたことを示します。

重複レコードタスクのルールおよびガイドライン

データセット内の例外レコードを確認して更新するプロセスは共同プロセスです。前のタスクで別のユーザーが分析したクラスタに対して、作業を行うことができます。また、タスクの完了時に、作業したクラスタを別のユーザーに渡すこともできます。各ユーザーは、他のユーザーの作業を確認および更新できます。

重複レコードタスクの作業を行うときは、次に挙げるルールおよびガイドラインを考慮してください。

- クラスタは、類似するまたは同一のデータ値を共有する、データベーステーブル内のレコードのセットです。レコードをクラスタにソート処理する条件を開発者が定義します。レコードが現在のクラスタに属していないと思われる場合、[検出されたクラスタ] オプションを使用して正しいクラスタを探します。現在のクラスタに属しているレコードが別のクラスタに含まれていると思われる場合、[検出されたクラスタ] オプションを使用して正しいクラスタを探します。
- クラスタに対する作業を行うときに、優先レコードを使用して、クラスタで表される最も正確なバージョンのビジネスエンティティを定義します。優先レコードは必ずしもレコードの最終バージョンである必要はありません。タスクが完了したら、別のユーザーまたは別のデータ処理がクラスタに対する作業を実行する場合があります。

優先レコードを更新するときに、クラスタ内のレコードの優先フォームを表す、例外データベース内のレコードを更新します。クラスタ内のソースデータベースは更新しません。

- クラスタのステータスを設定したら、クラスタを確認したことを示すことができます。この確認ステータスは、クラスタ内の優先レコードの正確性やデータ品質を説明するものではありません。
ベストプラクティスとして、調べたすべてのクラスタを確認済みとしてマークすることをお勧めします。ダウストリームタスクのユーザーは、別のユーザーがクラスタを調べたことをこのステータスで確認できます。レコード内のデータを更新するときは、レコードに対する別のステータスインジケータの有無に関係なく、レコードを確認済みとしてマークします。
- 監査証跡には、優先レコードにユーザーが加えたすべての変更が保存されます。監査証跡に、クラスタデータの変更は保存されません。
- 作業したデータは、データの修正タスクやデータの確認タスクに渡すことができます。例えば、ワークフロー内のヒューマンタスクを設定する開発者は、複数の例外修正タスクを順番に指定することができます。開発者は重複修正タスクと共に重複確認タスクを実行することができます。

第 4 章

タスク管理

この章では、以下の項目について説明します。

- [タスク管理の概要, 44 ページ](#)
- [タスク実行者の操作, 44 ページ](#)
- [ビジネス管理者の操作, 46 ページ](#)
- [タスクデータのエクスポート, 49 ページ](#)

タスク管理の概要

ワークフローのヒューマンタスクの設定を担当する開発者は、タスクデータを操作できるユーザーを指名します。Analyst ツールにログインすると、[開始] ワークスペースに、ワークフローによってログインユーザーに割り当てられているタスクインスタンスが一覧表示されます。ユーザーはタスクの実行者またはビジネス管理者としてタスクを操作できます。

タスク実行者は、タスク内のレコードおよびステータスインジケータを更新できます。ビジネス管理者は、タスク実行者と同じ操作を実行できます。それに加えて、タスクの別ユーザーへの再割り当てやワークフロー内の次のステージへのタスクの移動を実行することもできます。

ワークフローでは、ユーザーに、タスクごとに異なるロールを割り当てることができます。ユーザーは、1つのワークフロー内の異なるタスクで、タスク実行者になることもあればビジネス管理者になることもあります。

タスク実行者の操作

ワークフロー開発者は、ヒューマンタスクのインスタンスに対する作業を実行できるタスク実行者を選択します。自分が操作権限を保有しているタスクインスタンスの一覧を表示するには、[マイタスク] ビューを開きます。

[マイタスク] ビューでは、次の操作を実行できます。

タスクの内容を確認する

タスク内のレコードまたはクラスタを確認します。タスクの内容は、**[例外]** ワークスペースに表示されます。

タスクを開く

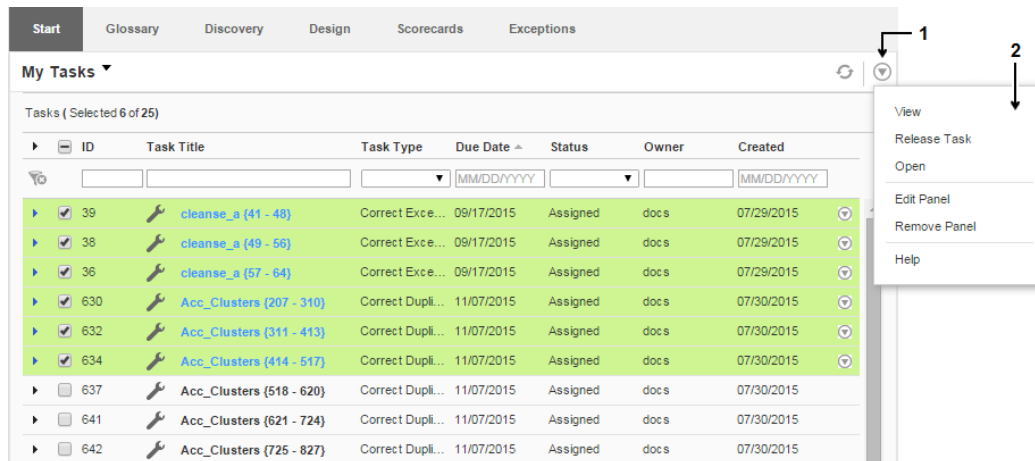
タスクを開いてレコードを確認および編集します。タスクを開いたユーザーが、そのタスクの所有者になります。その他のタスク実行者がそのタスクを開くことはできません。

タスクを解放する

自分が所有しているタスクを解放します。解放されたタスクは、ビジネス管理者が新しい所有者を割り当てるか、別のユーザーがそのタスクを開くまで、所有者なしのままになります。

タスクはいつでも解放できます。タスクを解放するためにそのタスクに対する操作を完了する必要はありません。

次の画像は、[開始] ワークスペースの [マイタスク] ビューを示しています。



タスクを実行するには、次の2つの方法があります。

1. [タスクメニュー] アイコン。
2. [タスクメニュー] オプション。

一度に複数のタスクに対して、表示、開く、解放の各操作を実行できます。

タスクの表示

タスクを表示すると、読み取り専用モードでタスクが開きます。タスク実行者は、所有者のいないタスクの内容を表示できます。ビジネス管理者は、[タスク管理] ビューの任意のタスクの内容を表示できます。

一度に複数のタスクを読み取り専用モードで開くことができます。各タスクは、[例外] ワークスペースで別々のタブに開きます。タスクチェックボックスを使用してタスクを選択します。

1. [開始] ワークスペースで、[マイタスク] ビューを開きます。
2. 1つ以上のタスクを選択します。

注: タスク名をクリックしないでください。タスク名をクリックすると、タスクを開いて、タスクの所有権を要求することになります。

3. [タスク] メニューから **[表示]** を選択します。

選択したタスクが、[例外] ワークスペースに開きます。

タスクを開く

タスクデータを操作するためにタスクを開きます。タスクを開いて、タスクの所有権を要求します。あなたが所有しているタスクを別のタスク実行者が編集することはできません。

一度に複数のタスクを開くことができます。各タスクは、[例外] ワークスペースで別々のタブに開きます。タスクチェックボックスを使用してタスクを選択します。

1. [開始] ワークスペースで、[マイタスク] ビューを開きます。

2. 1つのタスクを開く場合は、タスク名をクリックします。
または、1つ以上のタスクを選択します。
3. [タスク] メニューから **【開く】** を選択します。
選択したタスクが、[例外] ワークスペースに開きます。

タスクの解放

タスクを解放すると、そのタスクの所有者ではなくなります。別のユーザーがそのタスクの所有権を要求できるようになります。または、ビジネス管理者がそのタスクを別のユーザーに割り当てることができるようになります。タスクには、タスクデータに対して実行された操作がすべて保存されます。

一度に複数のタスクを解放できます。タスクチェックボックスを使用してタスクを選択します。ビジネス管理者は、[タスク管理] ビューで任意のタスクを解放できます。

1. [開始] ワークスペースで、[マイタスク] ビューを開きます。
2. 1つ以上のタスクを選択します。
3. [タスク] メニューから **【タスクの解放】** を選択します。

Analyst ツールによってタスクの一覧が更新され、当該タスクが所有者なしの状態になったことが示されます。

ビジネス管理者の操作

ワークフロー開発者から、ワークフローが生成するタスクインスタンスのビジネス管理者として指名を受ける可能性があります。自分が管理者となっているタスクインスタンスの一覧を表示するには、[タスク管理] ビューを開きます。

ビジネス管理者は、タスクのユーザーへの割り当て、ユーザーが実行する操作の表示、ワークフローの次のステージへのタスクの移動を実行できます。また、タスク実行者と同じように、タスクの所有権の要求やタスクデータの更新を実行することもできます。[タスク管理] ビューでは、タスク管理者が管理および所有しているタスクの一覧が表示されます。

[タスク管理] ビューでは、次の操作を実行できます。

タスクを表示するか、自分が所有しているタスクを開く

任意のタスクのデータを表示したり、自分が所有しているタスクを開きます。所有者のいないタスクの所有権を取得します。

タスクを再割り当てする

現在のユーザーがタスクをスケジュールどおりに完了できない場合に、別のユーザーにタスクを再割り当てします。

ユーザーがタスクで実行した操作を確認する

タスクサマリーデータを使用して、タスクでのユーザー作業の進行速度を測定します。

共通の親タスクを持つタスクの一覧を表示する

選択したタスクと同じヒューマンタスクを親に持つタスクの一覧を表示します。

一度に複数のタスクを完了する

選択したタスクと同じヒューマンタスクを親に持つすべてのタスクを完了します。タスクが完了すると、タスクデータがワークフローの次のステージに移動します。

一度に複数のタスクを表示および再割り当てできます。

ユーザーへのタスクの割り当て

所有者のいないタスクを割り当てたり、現在のユーザーがタスクを完了できない場合に別のユーザーまたはグループにタスクを割り当てます。

一度に複数のタスクを割り当てることができます。タスクチェックボックスを使用してタスクを選択します。

1. **【タスク管理】** ビューを開きます。
2. タスクリストから1つ以上のタスクを選択します。
3. **【タスク】** メニューから **【タスクの再割り当て】** を選択します。
4. ユーザーまたはグループの名前を選択します。
フィルタを使用してユーザーまたはグループを検索します。
5. 必要に応じて、コメントを入力します。
6. **【割り当て】** をクリックします。

【タスク管理】ビューでレコードが更新され、選択したタスク所有者が表示されます。

ヒューマンタスクでのタスクインスタンスのリストの表示

ワークフローが特定のヒューマンタスク用に生成したタスクインスタンスのリストを表示できます。

注: タスクインスタンスのリストを確認する際に、**【OK】** をクリックしないでください。 **【OK】** をクリックすると、タスクによって指定されたデータがワークフローの次のステージに移動します。

1. **【タスク管理】** ビューを開きます。
2. タスクリストからタスクを選択します。
3. **【タスク】** メニューから **【リンクされたタスクの完了】** を選択します。
4. タスクのリストを確認します。

タスクリストには、タスクごとに次の情報が表示されます。

- タスク ID。
ヒューマンタスク内のタスクインスタンスの一意的 ID です。
- タスクのタイトル。
【タスク管理】ビューに表示されるタスクの名前です。
- タスクのタイプ。
ヒューマンタスクが生成したタスクのタイプです。タスクのタイプとして、例外修正タスク、例外確認タスク、重複修正タスク、重複確認タスクがあります。
- タスクの所有者
タスクを所有しているユーザーの名前です。
- 期限
ユーザーがタスクを完了しなければならない期日です。
- ステータス
タスクの所有権のステータスです。ユーザーがタスクの所有権を要求する前は、タスクのステータスは **【作成済み】** になっています。ユーザーがタスクの所有権を要求するか、ユーザーにタスクが割り当てられると、タスクのステータスは **【割り当て済み】** になります。
- 作成日。
タスクを作成したワークフローが実行された日付です。

5. **【キャンセル】** をクリックします。

タスクデータをワークフローの次のステージに移動する場合は、**【OK】** をクリックします。

タスクサマリデータ

[例外] ワークスペースで表示するか開いた任意のタスクについて、サマリデータを読み取ることができます。サマリデータを使用すると、タスクでのユーザー作業の進行状況を測定できます。

重複レコードタスクには、タスク内のクラスタ数および確認されているクラスタの数が表示されます。例外レコードタスクには、タスク内のレコード数およびステータスインジケータを含むレコードの数が表示されます。

次の画像は、例外レコードタスクのサマリデータを示しています。

Start Glossary Discovery Design Scorecards Exceptions X

cleanse_a_rev (1 - 8) Save Cancel Undo Redo

To review exceptions, verify the work that a user performed in an earlier task. Review the data in each record, and review the record status. You can accept or undo the work of the user. When you finish the review of the records, update the task status.

8 of 8 Records are reviewed 6 Accepted 2 Reprocessed 0 Rejected

1 2 3 4

Data Editing Data Audit Find & Replace Filter

			STRING_COL	INT_COL	BIGINT_COL	NUMBER_COL	NSTRING_COL	DOUBLE_COL	DATETIME_COL
1			aaaaaaaa	11,111	8,890,000,000,000,0...	50	bbbbbb	7.527	01/01/2015 12:00:00
2			aaaaaaaa	11,111	8,890,000,000,000,0...	50	bbbbbb	7.527	01/02/2015 12:00:00

タスクには、次のサマリデータ値が格納されています。

1. ユーザーが確認したレコード数およびタスク内の合計レコード数。
2. ビジネスデータとして格納する有効なレコードとしてユーザーが承認したレコードの数。
3. 詳細な分析対象としてユーザーが選択したレコードの数。
4. ビジネスデータとして格納しない無効なレコードとしてユーザーが拒否したレコードの数。

タスク内のユーザー作業の進行状況の確認

タスクでのユーザー作業の進行状況を測定するには、サマリデータを確認します。

1. **【開始】** ワークスペースで、**【タスク管理】** ビューを開きます。
2. タスクを選択します。
[例外] ワークスペースにタスクが開きます。
3. タスクのサマリデータを確認します。

複数タスクの完了

1つのヒューマンタスクから生成されたすべてのタスクインスタンスを識別して、それらのタスクインスタンスを1回の操作で完了できます。

特定のヒューマンタスクのすべてのタスクインスタンスが完了すると、各タスクのレコードが、ワークフローの次のステージに移動します。Analyst ツールにより、タスクを操作したユーザーからタスクが削除されます。タスクの完了時に、レコードデータやステータスデータを更新する必要はありません。

次のような場合に、タスクを完了します。

- ワークフローに失敗し、ワークフローを再度実行する場合。
- ユーザーがタスクインスタンスを時間どおりに終了できない場合。

複数タスクの完了

特定のヒューマンタスクのすべてのタスクインスタンスを完了するには、そのヒューマンタスクから任意のタスクインスタンスを選択します。

1. [タスク管理] ビューを開きます。
2. タスクリストからタスクを選択します。
3. [タスク] メニューから **[リンクされたタスクの完了]** を選択します。

[リンクされたタスクの完了] ダイアログボックスに、選択したタスクと同じヒューマンタスクから生成されたタスクインスタンスのリストが表示されます。

4. 必要に応じて、コメントを入力します。
5. **[OK]** をクリックします。

タスクが完了します。タスクデータがワークフローの次のステージに移動します。

タスク完了後に受信トレイを開くと、タスクリストへの変更が受信トレイに表示されないことがあります。現在のタスクのリストを受信トレイで表示するには、受信トレイを更新します。

タスクデータのエクスポート

タスクインスタンスから区切りファイルにデータをエクスポートできます。データをエクスポートすると、データの現在の状態を他のユーザーと共有できます。

タスクをエクスポートすると、レコードデータ、ステータスインジケータデータ、およびワークフロー識別子の値がエクスポートされます。エクスポート操作を実行すると、ステータスインジケータデータおよびワークフロー識別子の値のためのカラムが追加されます。監査証跡データは、エクスポート操作の対象から除外されます。

例外タスクのメタデータ

例外修正タスクまたは例外確認タスクからデータをエクスポートすると、すべてのタスクデータとメタデータがエクスポートされます。エクスポート操作を実行すると、区切りファイルの先頭にメタデータカラムが追加されます。エクスポート操作により、例外修正タスク用と例外確認タスク用に同じカラムが作成されます。

次の表では、タスクデータとともにエクスポートするメタデータカラムについて説明します。

カラム名	説明
ROW_IDENTIFIER	データベーステーブル内のレコード行の数。
REVIEW_STATUS	レコードの現在の確認ステータスです。 レコードの確認ステータス値は、以下のいずれかになります。 <ul style="list-style-type: none">- NULL:ユーザーが確認ステータスを更新しませんでした。- REVIEWED:ユーザーがレコードを確認しました。- REJECTED:ユーザーがレコードを拒否しました。- ACCEPTED:ユーザーがレコードを承認しました。

カラム名	説明
WORKFLOW_ID	タスクインスタンスを生成したワークフローの識別子。
USER_COMMENT	ユーザーが Analyst ツールでレコードに追加した最新のメモ。
UPDATED_STATUS	レコードデータの現在のステータス。 レコードのステータス値は、以下のいずれかになります。 - UPDATED:ユーザーがレコードを更新しました。 - ACCEPTED:ユーザーがレコードを有効として承認しました。 - REJECTED:ユーザーがレコードを無効として拒否しました。 - REPROCESS:ユーザーが、レコードに別のユーザーまたは別のデータ処理による詳細な分析が必要であることを示しました。 - NULL:ユーザーがレコードステータスを更新しませんでした。
RECORD_STATUS	ワークフローで設定するレコードのステータス。ワークフローは、例外データテーブルにレコードを書き込むとき、ステータス値を設定します。デフォルトのステータスは INVALID です。

重複タスクのメタデータ

重複修正タスクまたは重複確認タスクからデータをエクスポートすると、すべてのタスクデータとメタデータがエクスポートされます。エクスポート操作を実行すると、区切りファイルの先頭にメタデータカラムが追加されます。エクスポート操作により、重複修正タスク用と重複確認タスク用に同じカラムが作成されます。

次の表では、タスクデータとともにエクスポートするメタデータカラムについて説明します。

カラム名	説明
ROW_IDENTIFIER	データベーステーブル内のレコード行の数。
SEQUENTIAL_CLUSTER_ID	データベーステーブル内のクラスタに対する識別子の値です。ワークフローは、この値を使用してデータベース内のクラスタ行をソートします。
CLUSTER_ID	レコードが属するクラスタを識別する値です。
MATCH_SCORE	クラスタ内の 2 つのレコードがどの程度似ているかを示す値です。スコアは、0 と 1 の間の小数値です。
IS_MASTER	テーブル内の優先レコードを識別する値です。優先レコードの場合は Y、その他のレコードの場合は N です。
UPDATED_STATUS	クラスタ内のレコードの更新ステータスです。 レコードのステータス値は、以下のいずれかになります。 - UPDATED:ユーザーがレコードを更新しました。 ユーザーは、クラスタ内の優先レコードを更新できます。 - NULL:ユーザーがレコードを更新しませんでした。 - EXTRACTED:ユーザーが元のクラスタからレコードを移動しました。
USER_COMMENT	ユーザーが Analyst ツールでクラスタに追加した最新のメモ。

カラム名	説明
REVIEW_STATUS	<p>クラスタの現在の確認ステータスです。</p> <p>クラスタの確認ステータス値は、以下のいずれかになります。</p> <ul style="list-style-type: none"> - NULL: ユーザーが確認ステータスを更新しませんでした。 - REVIEWED: ユーザーが、レコードが格納されているクラスタを確認しました。 - REJECTED: ユーザーが、レコードが格納されているクラスタを拒否しました。 - ACCEPTED: ユーザーが、レコードが格納されているクラスタを承認しました。
WORKFLOW_ID	タスクインスタンスを生成したワークフローの識別子。

タスクデータのエクスポート

タスクデータを区切りファイルにエクスポートします。

1. タスクを開きます。
2. [タスクアクション] メニューから、**[データのエクスポート]** を選択します。
[エクスポート] ダイアログボックスが表示されます。
3. 必要に応じて、エクスポートファイル名を変更します。
デフォルトでは、ファイル名はタスク名と同じです。
4. カラム名をエクスポートデータの最初の行としてエクスポートするオプションを選択またはクリアします。
5. **[エクスポート]** をクリックします。

Analyst ツールによって、データがディレクトリ構造にエクスポートされます。

索引

I

Informatica Analyst インタフェース
ログイン [17](#)

え

エクスポートファイル
重複タスク [50](#)
例外タスク [49](#)

か

〔開始〕 ワークスペース
カラム [13](#)
確認タスク
クラスタを確認する手順 [41](#)
重複確認 [42](#)
監査証跡
重複レコードフィルタ [40](#)
例外レコードフィルタ [26](#)

く

クラスタ
クラスタデータのフィルタリング [40](#)
作成 [39](#)
ステータスの更新中 [37](#)
編集 [35](#)
メモの追加 [39](#)
レコードの検索 [37](#)
クラスタのステータスをクリア
説明 [37](#)
クラスタのフィルタリング
手順 [41](#)

す

ステータスインジケータ
重複レコードタスク [36](#)
例外タスク [22](#)

た

タスク
タスクインスタンス [10](#)
タスクデータのエクスポート [49, 51](#)
タスクとワークフロー [10](#)
タスクの解放 [46](#)
タスクの種類 [9](#)
重複確認 [33](#)

タスク (続く)

重複修正 [31, 32](#)
ヒューマンタスク [10](#)
ヒューマンタスクのステップ [11](#)
表示 [45](#)
開く [45](#)
マッピングタスク [10](#)
例外確認 [19](#)
例外修正 [19](#)
タスクインスタンス
定義 [10](#)
タスク管理
ユーザーへのタスクの割り当て [47](#)
タスク管理オプション [46](#)
タスクのエクスポート
説明 [49](#)
タスクの実行者
〔マイタスク〕 ビュー [44](#)

ち

重複タスク
エクスポートファイルの構造 [50](#)
重複レコード
確認ステップ [32](#)
クラスタステータスの更新 [37](#)
クラスタの作成 [39](#)
クラスタの編集 [35](#)
〔検出されたクラスタ〕 オプション [37](#)
修正ステップ [32](#)
重複確認タスク [33](#)
重複修正タスク [31, 32](#)

て

〔データ監査〕 タブ
重複レコードフィルタ [40](#)
例外レコードフィルタ [26](#)
データの値の検索と置換
例外レコード [27, 28](#)
〔データ編集〕 タブ
クラスタデータのフィルタリング [40](#)
重複レコードタスク [33](#)
〔データ編集〕 パネル
クラスタのフィルタリング [41](#)
例外データフィルタ [25](#)
レコードのフィルタリング [26](#)

ひ

ビジネス管理者 [11, 44](#)

ふ

フィルタ
 クラスタデータ [40](#)
フィルタのオプション
 例外データ [24](#)
プロセスフロー
 重複レコードの例 [16](#)
 不良レコードの例 [15](#)

ま

「マイタスク」 ビュー [12](#), [44](#)

め

メモ
 クラスタへの追加 [39](#)

ゆ

優先レコード
 変更 [35](#)

れ

例
 重複レコードのプロセスフロー [16](#)
 不良レコードのプロセスフロー [15](#)
 例外タスクの検証エラー [22](#)

例外管理

 概要 [9](#)

例外タスク

 エクスポートファイルの構造 [49](#)
 検証エラー [22](#)

例外タスクの検証エラー [22](#)

例外レコード

 確認ステップ [18](#)
 修正ステップ [18](#)
 データの値の検索と置換 [27](#), [28](#)
 フィルタリング [26](#)
 例外確認タスク [19](#)
 例外修正タスク [19](#)
 例外レコードの編集 [21](#)
 レコードステータスの更新 [23](#)

「例外」 ワークスペース

 Informatica Analyst [14](#)
 「データ編集」 タブ [20](#)
 例外修正タスク [20](#)

ろ

ロール

 タスクの実行者 [11](#), [44](#)
 ビジネス管理者 [11](#), [44](#)

わ

ワークフロー

 説明 [10](#)