



Copyright © 2016 NTT DATA INTRAMART CORPORATION

目次

- 1. 改訂情報
- 2. はじめに
 - 2.1. 本書の目的
 - 2.2. 対象読者
 - 2.3. 本書の構成
- 3. intra-mart Accel GroupMail の全文検索機能
 - 3.1. Accel GroupMail クローラ
 - 3.2. 処理進捗の管理方式について
 - 3.3. コンテンツの情報
 - 3.4. ファセット
 - 3.5. 閲覧可能権限
 - 3.6. 検索結果表示項目

変更年月日	変更内容
2016-12-01	初版
2017-08-01	第2版 以下を追加・変更しました <ul style="list-style-type: none">▪ <i>intra-mart Accel GroupMail</i> の全文検索機能を更新しました
2022-12-01	第3版 以下を追加・変更しました <ul style="list-style-type: none">▪ <i>intra-mart Accel GroupMail</i> の全文検索機能に 実行パラメータ「saveOnSentToSolr」についての説明を追加しました

はじめに

項目

- 本書の目的
- 対象読者
- 本書の構成

本書の目的

本書では intra-mart Accel GroupMail の各機能の全文検索の機能概要と仕組の詳細について説明します。

対象読者

次の利用者を対象としています。

- intra-mart Accel GroupMail の各機能の全文検索機能（IM-ContentsSearch）を管理する運用担当者

次の内容を理解していることが必須となります。

- intra-mart Accel Platform の仕様
- IM-ContentsSearch の仕様
- intra-mart Accel GroupMail の仕様

本書の構成

本書は次の構成となっています。

- *intra-mart Accel GroupMail の全文検索機能*
intra-mart Accel GroupMail の全文検索機能の共通的な仕様について説明します。



コラム

用語解説

Storage として使用するディレクトリを %PUBLIC_STORAGE_PATH% と略します。
「intra-mart Accel Platform」 以下、iAP と略します。
「intra-mart Accel GroupMail」 以下、iAG と略します。

intra-mart Accel GroupMail の全文検索機能

この章では、iAG の全文検索機能における共通的な仕様について説明します。

項目

- Accel GroupMail クローラ
 - 差分クローリング
 - 削除クローリング
 - 再作成クローリング
- 処理進捗の管理方式について
 - 最終クロール日時による管理
 - 最終処理済みメール更新日時による管理
 - 処理進捗保存ファイルについて
- コンテンツの情報
 - 標準フィールド
 - 動的フィールド
- ファセット
- 閲覧可能権限
 - 全文検索時の閲覧権限
 - 各機能の情報に対する権限
- 検索結果表示項目

Accel GroupMail クローラ

iAG のコンテンツを作成するためのクローラはジョブスケジューラ機能のジョブとして提供しています。

各機能のクローラのジョブはテナント環境セットアップ時に「Accel GroupMail クローラ」ジョブカテゴリに登録されます。

ジョブカテゴリID iag-crawler-job-category

カテゴリ名	Accel GroupMail クローラ
-------	----------------------

クローラのジョブは標準で以下のジョブネットにより登録されています。

- 差分クローリング（ジョブネットID: crawler-delta-jobnet）
- 削除クローリング（ジョブネットID: crawler-delete-jobnet）
- 再作成クローリング（ジョブネットID: crawler-reindex-jobnet）

差分クローリング

差分クローリングのジョブネットが実行されたとき、前回のクローリング処理から新規に送受信、保存

したメールに対して差分クローリングが実行されます。

変更したコンテンツについては、以下の操作を実行した場合に差分クローリングの対象となります。

- メールを送信した場合
- メールを保存した場合
- メールを受信した場合

差分クローリング実行時の実行パラメータ

差分クローリング実行時には差分クローリングのジョブネット及びジョブに設定されている以下の実行パラメータが利用されます。

キー	説明
crawlingType	クローラの動作タイプを設定します。"DELTA"である場合のみ差分クローリングが動作します。
groupName	処理対象とする検索サーバグループIDを指定します。
maxSegments	最適化の度合いを設定します。
updateCount	Solrサーバへのコミット1回あたりに登録するコンテンツの件数
withCommit	処理終了時に更新データを確定するかを設定します。
withOptimize	処理終了時にデータの最適化をするかを設定します。
saveOnSentToSolr	インデックス化対象とするメールの処理済みのデータの管理方式を変更します。（詳細については後述の処理進捗の管理方式についてを参照） true : 処理進捗の管理方式として「最終処理済みメール更新日時による管理」を使用します。 false : (デフォルト) 処理進捗の管理方式として「最終クロール日時による管理」を使用します。



コラム

正常終了後はジョブ実行結果に以下のログが出力されます。

```
[SUCCESS] Accel GroupMail クローラ 処理対象メール: {0} 件、 処理時間: {1} msec  
( {0} : コンテンツ登録件数、 {1} : 差分クローリングの処理時間 )
```



注意

updateCountの数値が大きい場合、Solrサーバへのコミット1回あたりに登録するコンテンツの件数が増加するため、Solrサーバへの負荷が高くなりパフォーマンスに影響があります。

削除クローリング

削除クローリングのジョブネットが実行されたとき、全文検索機能（IM-ContentsSearch）に登録さ

削除クローリング実行時の実行パラメータ

削除クローリング実行時には削除クローリングのジョブネット及びジョブに設定されている以下の実行パラメータが利用されます。

キー	説明
crawlingType	クローラの動作タイプを設定します。"DELETE"である場合のみ削除処理は動作します。
groupName	処理対象とする検索サーバグループIDを指定します。
maxSegments	最適化の度合いを設定します。
withCommit	処理終了時に更新データを確定するかを設定します。
withOptimize	処理終了時にデータの最適化をするかを設定します。
saveOnSentToSolr	インデックス化対象とするメールの処理済みのデータの管理方式を変更します。（詳細については後述の 処理進捗の管理方式について を参照） true : 処理進捗の管理方式として「最終処理済みメール更新日時による管理」を使用します。 false : （デフォルト）処理進捗の管理方式として「最終クロール日時による管理」を使用します。

再作成クローリング

再作成クローリングのジョブネットが実行されたとき、iAG で登録した全てのコンテンツに対して、以下の手順でクローラの再作成処理が実行されます。

- 全文検索機能（IM-ContentsSearch）に登録されている iAG の全てのコンテンツ情報を削除
- iAG のコンテンツ情報の再登録

再作成クローリング実行時の実行パラメータ

再作成クローリング実行時には再作成クローリングのジョブネット及びジョブに設定されている以下の実行パラメータが利用されます。

キー	説明
crawlingType	クローラの動作タイプを設定します。"REINDEX"である場合のみ再作成クローリングが動作します。
groupName	処理対象とする検索サーバグループIDを指定します。
maxSegments	最適化の度合いを設定します。
updateCount	Solrサーバへのコミット1回あたりに登録するコンテンツの件数
withCommit	処理終了時に更新データを確定するかを設定します。

キー	説明
withOptimize	処理終了時にデータの最適化をするかを設定します。
saveOnSentToSolr	インデックス化対象とするメールの処理済みのデータの管理方式を変更します。（詳細については後述の 処理進捗の管理方式について を参照） true : 処理進捗の管理方式として「最終処理済みメール更新日時による管理」を使用します。 false : (デフォルト) 処理進捗の管理方式として「最終クロール日時による管理」を使用します。



コラム

正常終了後はジョブ実行結果に以下のログが 출력されます。

[SUCCESS] Accel GroupMail クローラ 処理対象メール: {0} 件、 処理時間: {1} msec
({0} : コンテンツ登録件数、 {1} : 再作成クローリングの処理時間)



注意

updateCountの数値が大きい場合、Solrサーバへのコミット1回あたりに登録するコンテンツの件数が増加するため、Solrサーバへの負荷が高くなりパフォーマンスに影響があります。

処理進捗の管理方式について

実行パラメータによってどこまでインデックス化したかの管理方式を変更する事ができます。 管理方式として最終クロール日時と、最終処理済みメール更新日時の2種があります。

最終クロール日時による管理

最終クロール日時による管理の場合、処理が正常終了した場合にクローラの最終実行日時を処理進捗保存ファイルに保存します。

差分クローリング実行時にはこのファイルを参照し、前回の実行日時より後にコンテンツの登録・更新対象となったメールを処理対象とします。

updateCount 毎にSolrへコンテンツを送信しますが、最終実行日時を更新するのはその時対象となつたメールすべてを処理した後になります。

このため、途中でエラーが発生するなどした場合は最終実行日時を更新しないため、次回は同じ日時から再処理することになります。

最終処理済みメール更新日時による管理

最終処理データの更新日時による管理の場合、updateCount 毎にSolrへコンテンツを送信したタイミングで

その中にあったメールの更新日時のうち、最も遅いものを処理進捗保存ファイルに保存します。

差分クローリング実行時にはこのファイルを参照し、その日時以降にコンテンツの登録・更新対象と

なったメールを処理対象にします。

このため、途中でエラーが発生した場合でもエラー発生直前に保存した日時以降のメールを対象に処理を続行できます。

ただし、取りこぼし防止の為保存された日時（秒単位）と同一の時刻 "以降" を対象とするため、必ず前回処理済みの末尾を重複処理します。

Solrへのコンテンツの登録が特に問題なく、安定して進む場合は極力最終クロール日時による管理を使用することを推奨します。

処理進捗保存ファイルについて

処理進捗保存場所

- %PUBLIC_STORAGE_PATH%/products/im_contents_search/store/%テナントID%/last_crawling_date/iag.json

処理進捗保存ファイルの内容

- JSON形式で "yyyy-MM-dd HH:mm:ss" 形式の日付文字列を保存します。
削除クローリングが実行された場合、"2000-01-01 00:00:00" に初期化します。



コラム

初回実行時の動作仕様

初期状態の場合、最終クロール日時を保存するファイルは存在していません。

この場合、既定値として"2000年1月1日0時0分0秒"を最終クロール日時として動作する仕様になっています。

初回実行時の対象日時を明示的に指定したい場合は、上記ファイルを作成して対象日時を上記の日時フォーマットにて設定してください。

コンテンツの情報

Accel GroupMail クローラではメール情報を以下の内容で登録します。

標準フィールド

No	フィールド名	設定値	備考
1	ID	iag_SND_%メールID% iag_RCV_%メールID%	コンテンツを一意に識別するIDです。
2	TYPE	iag iag\$SAVE iag\$SEND iag\$RECEIVED	コンテンツのタイプです。
3	URL	メール詳細画面のURL	

No	フィールド名	設定値	備考
4	ID_ORIGINAL	メールID	コンテンツの元データにおける一意なキーとなるIDです。
5	TITLE	件名	コンテンツのタイトルです。
6	TEXT	本文（プレーンテキスト）	コンテンツの主となるテキストデータです。
7	ATTACHMENT	添付ファイル名 添付ファイル内のテキストデータ	
8	RECORD_DATE	メールの送信（保存）日時	

動的フィールド

No	フィールド名	設定値	備考
1	CONTENT_TYPE	SND RCV	コンテンツ種別を設定します。
2	REGISTER_USER_CD_STRING	登録者ユーザコード	差出人のユーザコードです。
3	REGISTER_USER_NAME_STRING	登録者ユーザ名	差出人のユーザ名です。
4	MAIL_SUB_ID	下書きメールのメールサブID	



コラム

タイトル（TITLE）、テキスト(TEXT)、添付ファイル(ATTACHMENT)に登録した内容が検索時に検索対象となるフィールドです。

ファセット

Accel GroupMail クローラではコンテンツ作成時に以下のファセット分類を適用しています。

GroupMail

Accel GroupMail クローラで作成されたすべてのコンテンツに対して "GroupMail" というファセットを適用します。

カテゴリ名

"GroupMail" のファセットを適用したコンテンツに対して、更にカテゴリ名のファセットを適用します。

- 下書きメールのコンテンツに対して"下書きメール"というファセットを適用します。

- 受信メールのコンテンツに対して”受信メール”というファセットを適用します。
- 送信済みメールのコンテンツに対して”送信済みメール”というファセットを適用します。



コラム

** Accel GroupMail クローラが全文検索画面に提供するファセットナビゲーションの例
**

全文検索画面の検索結果に iAG のコンテンツがある場合、以下の様なファセットナビゲーションが全文検索画面の左部に表示されます。
リンクを選択することで検索結果の絞込み検索を行うことができます。

絞り込み条件

- ▼ **GroupMail (48)**
 - 下書きメール (5)
 - 受信メール (28)
 - 送信済みメール (15)

閲覧可能権限

全文検索時の閲覧権限

クローラはジョブ実行日時の時点での権限を判定し、コンテンツに対して閲覧権限を適用します。
全文検索で閲覧可能なコンテンツは一般ユーザの閲覧権限により参照できる情報と同一です。

各機能の情報に対する権限

検索結果に表示されたコンテンツの画面表示・処理を行うことができるかどうかは一般ユーザ利用時の権限により判定されます。

検索結果表示項目

Accel GroupMail クローラが作成したコンテンツの検索結果で全文検索画面に表示されるコンテンツの表示項目は以下です。

- タイトル（件名）

メールの件名を表示します。

件名が未設定の場合は”<件名なし>”が表示されます。

- 登録日時

メールの送信（保存）日時を表示します。

- カテゴリ

メールのカテゴリを表示します。

- 受信メール GroupMail > 受信メール
- 送信済みメール GroupMail > 送信済みメール
- 下書きメール GroupMail > 下書きメール

- 差出人

メールを送信したユーザの名前を表示します。

- 要約

要約表示の指定をして検索した場合に表示されます。

要約は検索条件の文字列がカラーでハイライトされて表示されます。

検索結果画面例

新商品発表会開催のお知らせ
2016/06/09 14:23:30 GroupMail > 送信済みメール
差出人: 青柳辰巳
先日役員会議で決定した発表会の概要についてお知らせします。 新商品発表会 内容:新商品発表会 主催:サンプル会社 東京会場:2016年06月08日(月) 13:30~17:00 大阪会場:201

新商品発表会開催のお知らせ
2016/06/09 14:23:30 GroupMail > 受信メール
差出人: 青柳辰巳
先日役員会議で決定した発表会の概要についてお知らせします。 新商品発表会 内容:新商品発表会 主催:サンプル会社 東京会場:2016年06月08日(月) 13:30~17:00 大阪会場:201



コラム

検索結果のメールをクリックすると、新しいウィンドウで以下の画面を表示します。

■受信メールまたは送信済みメールの場合

選択したメールのメール詳細画面を表示します。

メール詳細情報

件名 新商品発表会開催のお知らせ
受信日時 2013/06/21 16:05
差出人 青柳辰巳
TO 円山益男, 林政義, 片山聰

先日役員会議で決定した発表会の概要についてお知らせします。

新商品発表会
内容:新商品発表会
主催:サンプル会社

東京会場:2013年07月01(月) 13:00~17:00
大阪会場:2013年07月02(火) 13:00~17:00
名古屋会場:2013年07月03(水) 13:00~17:00

■下書きメールの場合

選択したメールのメール作成画面を表示します。

メール作成

差出人 青柳辰巳
To: 円山益男, 林政義, 片山聰
件名 新商品発表会開催のお知らせ
重要 HTML

先日役員会議で決定した発表会の概要についてお知らせします。

新商品発表会
内容:新商品発表会
主催:サンプル会社

東京会場:2013年07月01(月) 13:00~17:00
大阪会場:2013年07月02(火) 13:00~17:00
名古屋会場:2013年07月03(水) 13:00~17:00

ファイル追加... 中断 削除