

# **IM-ContentsSearch for ISP**

## **Ver.7.0**

---

---

**ISP クローラ仕様書**

**2009/09/30 初版**



<< 変更履歴 >>

変更年月日	変更内容
2009/09/30	初版



## &lt;&lt; 目次 &gt;&gt;

1	はじめに.....	1
1.1	目的.....	1
1.2	前提条件.....	1
1.2.1	本資料の記載内容について.....	1
1.3	対象機能.....	1
2	システム概要.....	2
3	動作仕様.....	5
3.1	基本仕様.....	5
3.1.1	登録対象情報.....	5
3.1.2	アクセス権.....	7
3.1.3	更新・削除情報の反映.....	8
3.1.4	添付ファイル.....	8
3.1.5	最終クローラ起動日時の保存.....	10
3.1.6	クローラ実行時エラーの動作仕様.....	10
3.1.7	スタートバック区分コード管理情報.....	11
3.2	掲示板.....	12
3.2.1	クローラ.....	12
3.2.2	業務テンプレート表示情報.....	13
3.2.3	コンテンツ画面.....	14
3.3	電子会議室.....	14
3.3.1	クローラ.....	14
3.3.2	業務テンプレート表示情報.....	16
3.3.3	コンテンツ画面.....	16
3.4	ドキュメント管理.....	17
3.4.1	クローラ.....	17
3.4.2	業務テンプレート表示情報.....	18
3.4.3	コンテンツ画面.....	19
3.5	スケジュール.....	19
3.5.1	クローラ.....	19
3.5.2	業務テンプレート表示情報.....	21
3.5.3	コンテンツ画面.....	21
3.6	ブログ記事検索.....	22
3.6.1	クローラ.....	22
3.6.2	業務テンプレート表示情報.....	23
3.6.3	コンテンツ画面.....	23
3.7	営業日報.....	24
3.7.1	クローラ.....	24
3.7.2	業務テンプレート表示情報.....	25
3.7.3	コンテンツ画面.....	26
4	ISPクローラの拡張.....	27
4.1	リスナーの呼び出し.....	27
4.2	リスナの作成と設定.....	27
4.2.1	処理の記述.....	28



# 1 はじめに

## 1.1 目的

ISP クローラは、intra-mart と全文検索エンジンサーバ Solr (以下 Solr サーバと略します) との連携が行われている環境で、イントラネット・スタートパックの対象機能情報を収集し Solr サーバに文書として登録する機能を持つバッチプログラム群です。

本資料では、ISP クローラの動作仕様と拡張方法について解説します。

## 1.2 前提条件

- ISP クローラが動作するためには、
  - ・ IM-ContentsSearch for ISPの製品がインストールされている必要があります。  
セットアップ手順については、「IM-ContentsSearch for ISP セットアップガイド」をご参照ください。
- 本資料では全文検索機能固有の用語、イントラネット・スタートパック固有の用語が使用されています。その為、前提知識として全文検索機能とスタートパックの基本的な仕様を理解している必要があります。  
全文検索についての十分な知識を有していない場合は、「IM-ContentsSearch セットアップガイド」を、スタートパックについての十分な知識を有していない場合は、スタートパックのマニュアルを事前にお読みください。
- 本書に記述されているサンプルプログラムは、JavaEE 開発モデルで記述されています。そのため、JavaEE 開発モデルに関する理解は必須です。各開発モデルに関しては、付属する各種マニュアルおよび API リストを参照してください。

### 1.2.1 本資料の記載内容について

本資料では ISP クローラを拡張する際のソースコードサンプルを掲載しています。

サンプルコードでは拡張機能について理解することに主眼をおいていますので、必ずしもベストなコーディング方法とはいえない方法もあえて取っている箇所があります。あくまでも、サンプルとしての位置付けでとらえるようにしてください。

## 1.3 対象機能

イントラネット・スタートパックが提供する機能の中で、以下の機能が IM-ContentsSearch for ISP の対象機能です。

- 掲示板
- 電子会議室
- ドキュメント管理
- スケジュール
- 営業日報
- ブログ記事

## 2 システム概要

ISP クローラは、intra-mart 上で intra-mart Batch Server により実行されるバッチプログラム群です。

各バッチプログラムを起動すると、インデックス作成対象の情報を取得し、IM-ContentsSearch クライアントAPIを利用して、Solr サーバに取得した情報を登録します。

文書を登録し、インデックスが作成されたスタートパックの情報は、intra-martの全文検索画面からの検索が可能となります。

以下に ISP クローラのシステム構成を示します。

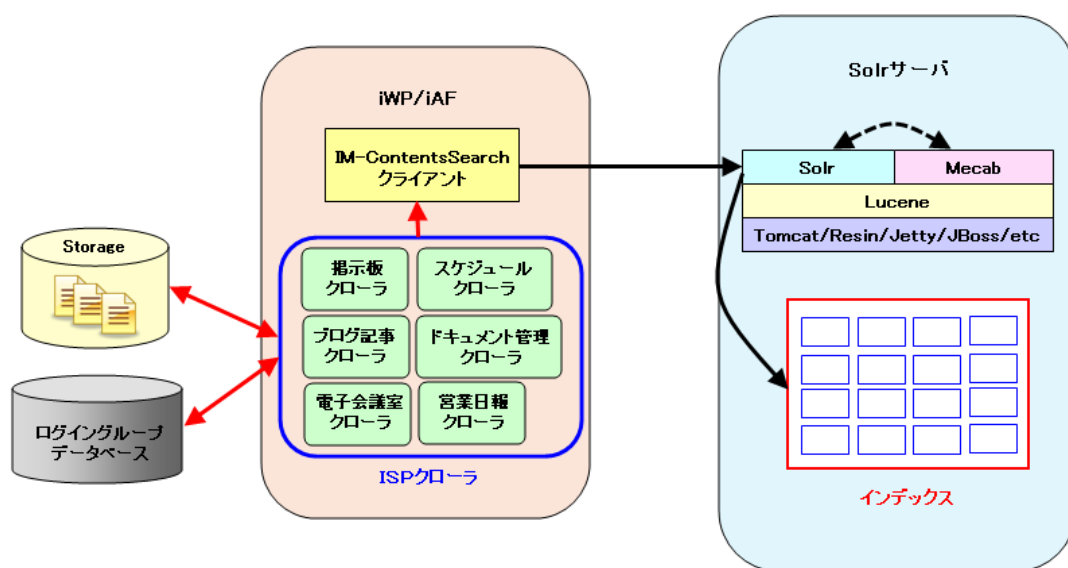


図 2-1 ISP クローラ システム概要

本機能では、使用用途に応じて各機能単位に 2 種類のクロウラを提供します。

- 差分クロウラ・・・最終クロウラ実行日時以降に登録または更新された情報を登録対象とします。
- 再作成クロウラ・・・差分クロウラの機能に加え、既存の文書を削除と対象情報の再登録を行います。

以下にバッチ登録情報を示します。

### 1. 掲示板

差分	バッチ名	掲示板 差分クロウラ
	バッチ ID	isp_cs_bllt_crwlr_d
	実行プログラム言語	JAVA
	実行プログラムパス	jp.co.intra_mart.product.startpack.contents_search.bulletin.BulletinDeltaCrawler
再作成	バッチ名	掲示板 再作成クロウラ
	バッチ ID	isp_cs_bllt_crwlr_r
	実行プログラム言語	JAVA
	実行プログラムパス	jp.co.intra_mart.product.startpack.contents_search.bulletin.BulletinReIndexingCrawler



## 2. 電子会議室

差分	バッチ名	電子会議室 差分クローラ
	バッチ ID	isp_cs_cnfr_cwlr_d
	実行プログラム言語	JAVA
	実行プログラムパス	jp.co.intra_mart.product.startpack.contents_search.conference. ConferenceDeltaCrawler
再作成	バッチ名	電子会議室 再作成クローラ
	バッチ ID	isp_cs_cnfr_cwlr_r
	実行プログラム言語	JAVA
	実行プログラムパス	jp.co.intra_mart.product.startpack.contents_search.conference. ConferenceReIndexingCrawler

## 3. ドキュメント管理

差分	バッチ名	ドキュメント管理 差分クローラ
	バッチ ID	isp_cs_dcmn_cwlr_d
	実行プログラム言語	JAVA
	実行プログラムパス	jp.co.intra_mart.product.startpack.contents_search.document. DocumentDeltaCrawler
再作成	バッチ名	ドキュメント管理 再作成クローラ
	バッチ ID	isp_cs_dcmn_cwlr_r
	実行プログラム言語	JAVA
	実行プログラムパス	jp.co.intra_mart.product.startpack.contents_search.document. DocumentReIndexingCrawler

## 4. スケジュール

差分	バッチ名	スケジュール 差分クローラ
	バッチ ID	isp_cs_schd_cwlr_d
	実行プログラム言語	JAVA
	実行プログラムパス	jp.co.intra_mart.product.startpack.contents_search.schedule. ScheduleDeltaCrawler
再作成	バッチ名	スケジュール 再作成クローラ
	バッチ ID	isp_cs_schd_cwlr_r
	実行プログラム言語	JAVA
	実行プログラムパス	jp.co.intra_mart.product.startpack.contents_search.schedule. ScheduleReIndexingCrawler

## 5. ブログ記事

差分	バッチ名	ブログ記事 差分クローラ
	バッチ ID	isp_cs_blog_cwlr_d
	実行プログラム言語	JAVA
	実行プログラムパス	jp.co.intra_mart.product.startpack.contents_search.blog.BlogDeltaCrawler
再作成	バッチ名	ブログ記事 再作成クローラ
	バッチ ID	isp_cs_blog_cwlr_r
	実行プログラム言語	JAVA
	実行プログラムパス	jp.co.intra_mart.product.startpack.contents_search.blog.BlogReIndexingCrawler

## 6. 営業日報

差分	バッチ名	営業日報 差分クローラ
	バッチ ID	isp_cs_daily_crwlr_d
	実行プログラム言語	JAVA
	実行プログラムパス	jp.co.intra_mart.product.startpack.contents_search.daily.DailyDeltaCrawler
再作成	バッチ名	営業日報 再作成クローラ
	バッチ ID	isp_cs_daily_crwlr_r
	実行プログラム言語	JAVA
	実行プログラムパス	jp.co.intra_mart.product.startpack.contents_search.daily.DailyReIndexingCrawler

表 2-1 バッチ登録情報

## 3 動作仕様

この章では、ISP クローラの動作仕様について詳しく解説します。

### 3.1 基本仕様

各機能共通の基本的な動作仕様について説明します。

#### 3.1.1 登録対象情報

##### 3.1.1.1 登録に関する前提条件

intra-mart WebPlatform／AppFramework で提供する全文検索機能は、業務間を横断的に検索するアプリケーション共通の検索機能です。

その為検索キーワードに指定された語句以外の一部のアプリケーションのみに有効な条件を検索条件に含めて検索することはできません。

よって、ISPクローラが Solr サーバに登録・インデックス作成対象とする情報は、クローラ実行日時点でユーザが参照することが可能である有効な情報のみとします。

##### 3.1.1.2 差分クローラが対象とする範囲

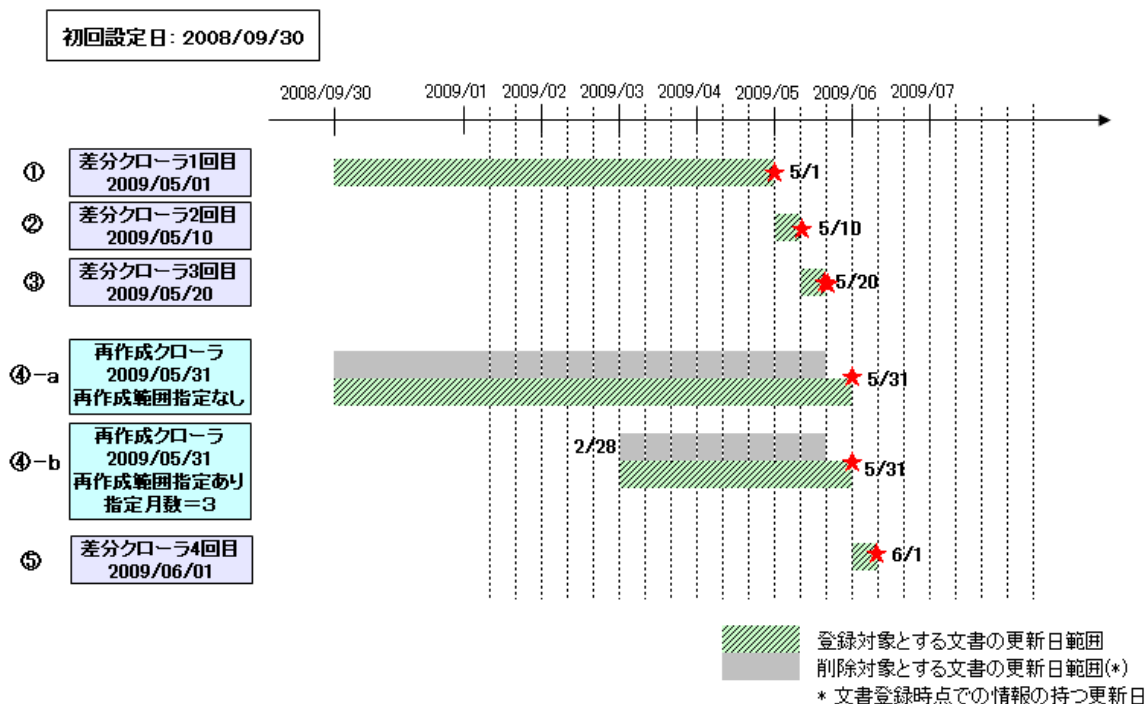
- 初回起動時(差分・再作成合わせた初回)  
区分コード※に設定されたデータ取得開始日(以降は**初回設定日**と略します)より後に登録・更新された情報を Solr サーバに登録します。
- 2回目以降  
最終クローラ起動日より後に登録・更新された情報を Solr サーバに登録します。

※区分コードについては、「3.1.7 スタートバック区分コード管理情報」をご参照ください。

##### 3.1.1.3 再作成クローラが対象とする範囲

- 再作成範囲指定なし・・・区分コードの再作成範囲指定月数=0の場合  
該当機能のインデックス情報を全て削除した後に、初回設定日より後に登録・更新された情報を Solr サーバに登録します。
- 再作成範囲指定あり・・・区分コードの再作成範囲指定月数 $\geq 1$ の場合
  - ◆ 初回起動時(差分・再作成クローラ合わせた初回)  
差分クローラの初回起動時と同じ
  - ◆ 2回目以降  
登録済インデックス情報の更新日が「クローラ起動日－指定月数」以降のものを削除した後、「クローラ起動日－指定月数」より後に登録・更新された情報を対象とします。  
但し、初回起動時または算出した指定日が初回設定日より前の場合は、初期設定値より後に登録・更新された情報を Solr サーバに登録します。

各クローラの実行例を以下に示します。



- ① 差分クローラ初回起動  
初回設定日 2008/09/30～クローラ起動日 2009/05/01 までに登録・更新された情報を登録
- ② 差分クローラ2回目  
前回クローラ起動日 2009/05/01～クローラ起動日 2009/05/10 までに登録・更新された情報を登録
- ③ 差分クローラ3回目  
前回クローラ起動日 2009/05/10～クローラ起動日 2009/05/20 までに登録・更新された情報を登録
- ④-a 再作成クローラ 範囲指定なし  
登録済みのインデックス情報を全て削除後、初回設定日 2008/09/30～クローラ起動日 2009/05/31 までに登録・更新された情報を登録
- ④-b 再作成クローラ 範囲指定 3ヶ月  
再作成日 2009/02/28 (2009/05/31－3ヶ月)以降のインデックス情報を削除後、再作成日～クローラ起動日 2009/05/31 までに登録・更新された情報を登録
- ⑤ 差分クローラ4回目  
前回クローラ起動日 2009/05/31～クローラ起動日 2009/06/01 までに登録・更新された情報を登録

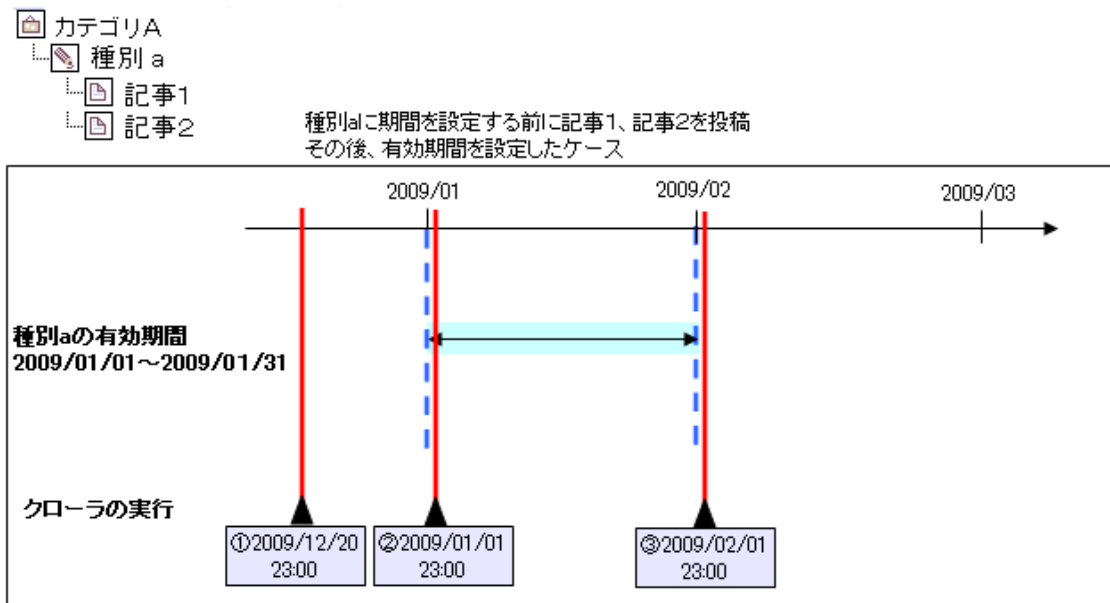
### 3.1.1.4 有効期間を持つ情報の登録・削除のタイミング

掲示板・電子会議室の種別は有効期間、また掲示板情報は掲載期間を持ち期間外には該当情報に紐付く情報を参照することはできません。

ISPクローラでは、クローラ実行日を基準にその情報の有効性を判断し、Solr サーバへ情報の登録・削除を行います。

※クローラを実行するまでは、期間を持つ情報に対する Solr サーバへの反映は行われませんのでご注意ください。

以下に、有効期間を持つ種別の情報が Solr サーバに登録・削除される例を示します。



① クローラ実行日【有効期間前】

「種別 a」は有効期間外のため、差分・再作成クローラ共に種別 a に紐付く「記事 1・2」を登録対象としません。

② クローラ実行日【有効期間前⇒中】

「種別 a」は有効期間内のため、差分・再作成クローラ共に種別 a に紐付く「記事 1・2」を登録対象とします。

③ クローラ実行日【有効期間中⇒後】

「種別 a」は有効期間外のため、差分・再作成クローラでは種別 a に紐付く「記事 1・2」のインデックス情報を削除し、以降は登録対象とはしません。

※クローラを実行する前は、期間外でも全文検索による「記事 1・2」の参照が可能です。

### 3.1.2 アクセス権

ユーザが全文検索で閲覧可能なスタートバック情報は、基本的にはユーザがスタートバックの画面により参照できる情報と同一ですが、以下の場合はスタートバック機能のアクセス権とは若干異なります。

■ スケジュール・営業日報

詳細は、3.5 スケジュール、3.7 営業日報のアクセス権をご参照ください

■ 掲載期限外の掲示板情報

・ スタートバックの機能からの参照時

投稿者・該当種別に管理権限が付与されたグループに所属するユーザが閲覧可能

・ 全文検索機能からの参照時

掲載期限後にクローラを実行した時点で、該当情報は削除されるため全てのユーザが閲覧不可

■ 掲示板・電子会議室・ドキュメント管理のアクセス権に役職による絞込み条件が設定されている場合

・ スタートバックの機能からの参照時

システム日付に該当組織に所属し、該当の役職を保持するユーザが閲覧可能

・ 全文検索機能からの参照時

クローラ実行日に該当組織に所属し、該当の役職を保持するユーザが閲覧可能

- コミュニティの掲示板・電子会議室・ドキュメント管理情報
    - ・ スタートバックの機能からの参照時  
システム日付に該当コミュニティに所属しているユーザが閲覧可能  
配信先に設定されたユーザ・グループに所属するユーザが配信期間内に閲覧可能
    - ・ 全文検索機能からの参照時  
クローラ実行日に該当コミュニティに所属しているユーザのみが閲覧可能  
配信先に設定されたユーザ・グループは対象外
- 《 コミュニティへの参加・脱退時におけるアクセス権の変更について 》
- Solr サーバに登録済の文書に、コミュニティのメンバ変更によるアクセス権の変更が反映されるのは、その情報自体に変更がなかった場合以下のタイミングになります。
- 新たにメンバが参加した場合  
差分・再作成クローラの実行時に反映
  - メンバが脱退した場合  
再作成クローラの実行時に反映
- 但し、範囲指定がある場合は、再作成範囲内の情報のみが対象となります。

### 3.1.3 更新・削除情報の反映

Solr サーバに情報を登録後、該当情報が更新・削除された場合に Solr サーバの登録情報に反映されるのは以下のタイミングになります。

- スタートバック関連情報が更新された場合  
差分・再作成クローラの実行時に反映  
【Solr サーバ未反映時の動作】  
文書登録時の情報で検索が行われ、検索結果画面に表示される。  
タイトルリンクを押下時に表示されるコンテンツ画面の情報は最新の情報が表示される。
- スタートバック関連情報が削除された場合  
再作成クローラの実行時に反映  
但し、範囲指定がある場合は、再作成範囲内の情報のみが対象となります。  
【Solr サーバ未反映時の動作】  
文書登録時の情報で検索が行われ、検索結果画面に表示される。  
タイトルリンクを押下時には、エラー画面が表示される。
- スタートバック以外の情報(所属組織など)が更新・削除された場合  
再作成クローラの実行時に反映  
但し、範囲指定がある場合は、再作成範囲内の情報のみが対象となります。  
【Solr サーバ未反映時の動作】  
文書登録時の情報で検索が行われ、検索結果画面に表示される。

### 3.1.4 添付ファイル

各機能の登録対象の情報に添付ファイルがある場合、ISP クローラでは添付ファイル情報もテキストフィールドに登録します。

登録される内容は以下の通りです。

- ◆ ドキュメントのファイルフォーマットがサポート対象外の場合
  - ・ ファイル名
- ◆ ドキュメントのファイルフォーマットがサポート対象の場合
  - ・ ファイル名
  - ・ ドキュメント内テキスト

#### 3.1.4.1 ドキュメント内テキストのインデックス化

ドキュメント内テキストがテキストフィールドに登録された場合は、ドキュメント内テキストのインデックスが作成され、全文検索の検索キーワードにドキュメントに含まれる単語を指定することで、該当情報を検索結果として取得することが可能となります。

但し一つのフィールドに索引付けする単語数は Solr サーバ側の設定に依存するため、サイズの大きいドキュメントや複数のドキュメントが添付ファイルとして登録されている場合、ドキュメントの最後までインデックス化されない可能性があります。

ドキュメント内のインデックス化されていない単語を検索キーワードに指定した場合には対象情報を取得することはできませんので、ご注意ください。

テキストを抽出する際の制限事項に関しては、「IM-ContentsSearch プログラミングガイド」 6.2.2 テキストを抽出する際の制限事項 をご参照ください。

#### 3.1.4.2 テキストを抽出することができるファイルフォーマット一覧

テキストを抽出できるファイルフォーマットの一覧を下表に示します。

但しファイルフォーマットとファイルの拡張子が下記表に一致しない場合は、サポート対象外とみなされます。

ファイルフォーマット	ファイルの拡張子
テキスト	txt
PDF	pdf
HTML	html,htm
XML	xml
ZIP	zip
Microsoft Word	doc, docx※
Microsoft Excel	xls, xlsx※
Microsoft PowerPoint	ppt, pptx※

表 3.2.7.2-1 サポート対象ファイルフォーマット一覧

※これらの拡張子のファイルは、テキストの抽出方法にオプションツールを使用する場合のみ対応しています。

詳細は「IM-ContentsSearch プログラミングガイド」 6.2.2 テキストを抽出する際の制限事項 をご参照ください。

#### 3.1.4.3 添付ファイルがサポート対象外の場合の動作仕様

サポート対象外のドキュメントが添付されていた場合は、警告ログを出力しファイル名のみインデックスを作成して処理を続けます。

#出力されるログの例

[WARN] j.c.n.i.s.SolrManager - 次のファイルは、テキスト抽出対象外です:。 [ファイル名]

### 3.1.4.4 添付ファイルからのドキュメント取り出しに失敗した場合の動作仕様

ドキュメントからのテキスト抽出に失敗した場合は、警告ログを出力しファイル名のみインデックスを作成して処理を継続します。

#出力されるログの例

```
[WARN] j.c.n.i.s.SolrManager - ファイルからテキストを抽出できません: [ファイルのパス]
```

### 3.1.4.5 パスワード付添付ファイルの場合の動作仕様

添付ファイルにパスワードが設定されている場合ドキュメントからテキストの抽出は行いません。  
警告ログを出力し、ファイル名のみインデックスを作成して処理を継続します。

またパスワード付zip ファイルの場合は、暗号化方式により動作が若干異なります。

#### ■ ZIP の暗号化方式が ZIP 2.0 互換の場合

zip ファイル名、ZIP 内のファイル名を登録し、警告ログを出力して処理を継続

#出力されるログの例

```
[WARN] j.c.n.i.s.u.ZipExtractor - ZIP 内の次のファイルは、暗号化されているため、処理されませんでした  
[Zip ファイル名]
```

#### ■ ZIP の暗号化方式が ZIP 2.0 互換以外 (AES-128bit・AES-256bit など) の場合

zip ファイル名を登録し、警告ログを出力して処理を継続

#出力されるログの例

```
[WARN] j.c.n.i.s.SolrManager - ファイルからテキストを抽出できません: [ファイルパス]
```

## 3.1.5 最終クローラ起動日時の保存

ISP クローラでは処理が正常終了した場合、クローラの最終起動日時を機能単位でファイルに保存します。

次回起動時にはこのファイルより前回のクローラ実行日を取得します。

最終起動日時保存ファイルは、クローラの初回正常終了時に以下のディレクトリ内に作成され、次回からは日時の更新のみが行われます。

日付のフォーマットは、yyyy/MM/dd|hh:mm:ss 形式です。

```
%StorageService%/ startpack/ contents_search/ cawler /%ログイングループ名%/
%機能名%_last_crawling_date
```

機能名一覧	
掲示板	bulletin
電子会議室	conference
ドキュメント管理	document
スケジュール	schedule
ブログ記事	blog
営業日報	daily

## 3.1.6 クローラ実行時エラーの動作仕様

クローラ実行中に予期しないエラーが発生してクローラが異常終了した場合、データベースとは異なり Solr サーバに対してロールバック処理を行い、インデックス情報をクローラ実行前の状態に戻すことは出来ません。

従って、今回のクローリングでエラーが発生する前に作成したインデックス情報はそのまま残ることになります。

但し、異常終了時には以下の処理は行いません。



- 登録済インデックスの即時反映
- インデックスの最適化
- 最終クローラ起動日の更新

#### 3.1.6.1 エラー発生時の作成済インデックス

クローラ実行中にエラーが発生して異常終了した場合、同一クローリングにてエラー発生前に作成したインデックスの即時反映処理は行われていないため、すぐに検索結果としてエラー発生前に登録した情報が取得されることはありません。しかし Solr サーバで何れかのタイミングで登録データのフラッシュが行われた場合、検索結果に反映される動作となります。

#### 3.1.6.2 エラー発生後の次回起動時の動作仕様

異常終了後、次にISP クローラを起動した時は、最後に正常終了した日時を前回のクローラ実行日として処理を行います。

前回のクローリングで登録済みの情報があつた場合、再び同一情報を登録することになりますが、Solr サーバでは登録時に同一IDの文書が既に存在した場合、文書を上書きする動作となっていますので再度登録処理を行っても問題はありません。

### 3.1.7 スタートパック区分コード管理情報

ISP クローラでは、スタートパックの区分コードとして以下の情報を管理しています。

これらの値を変更する際は、メニュー[マスタメンテナンス]－[区分コード設定] 「スタートパック共通」の[区分コード一覧]画面にて、該当項目の値を編集してください。

項番	情報項目	形式	キー	初期値
1	初回起動時のデータ取得 開始日(初回設定日) 《機能単位》	yyyy/MM/dd	crawling_date_initial_ /%機能名※%/	2008/09/30
2	再作成クローラの再作成範 囲指定月数 《機能単位》	Nヶ月前	re-indexing_limit_ /%機能名※%/	0

※機能名は 3.1.5 機能名一覧を参照

## 3.2 掲示板

### 3.2.1 クローラ

動的な文書種別情報の取得と掲示情報の Solr サーバへの登録・更新・削除を行います。

※動的な文書種別情報とは、全文検索結果画面の種別の絞込みに使用されるカテゴリ・種別情報です。

#### 3.2.1.1 登録・更新対象

<<基本条件>>

- ・クローラ実行日時時点で有効な掲示板種別・コミュニティ掲示板種別の掲示情報を対象とします。
- ・クローラ実行日時時点で掲載期間内の掲示情報を対象とします。

##### ■ 再作成クローラ(再作成範囲指定なし)

基本条件を満たした範囲期間内の全ての掲示情報を対象とします。

範囲期間に関しては、3.1.1.3 再作成クローラが対象とする範囲 に準じます。(以下同様)

##### ■ 差分クローラ、再作成クローラ(再作成範囲指定あり)

基本条件を満たし、かつ以下の条件のいずれかを満たす掲示情報を対象とします。

1. 前回のクローラ実行日または再作成日以降に掲示板種別が有効になった種別に紐付く更新日が初回設定日以降の掲示情報  
※再作成日・・・クローラ実行日－範囲指定月数 により算出された日付(以下同様)
2. 前回のクローラ実行日または再作成日以降にマスタ情報が更新されたカテゴリ・種別に紐付く更新日が初回設定日以降の掲示情報
3. 前回のクローラ実行日または再作成日以降に登録・更新された掲示情報
4. 前回のクローラ実行日または再作成日以降に掲載期間が開始した掲示情報
5. 前回のクローラ実行日または再作成日以降にコミュニティに参加したメンバがいる更新日が初期設定値以降のコミュニティの掲示情報

#### 3.2.1.2 削除対象

##### ■ 再作成クローラ

範囲期間内の掲示板に紐付く全ての文書を削除します。

##### ■ 差分クローラ

1. 前回のクローラ実行日以降に有効期限切れになった掲示板種別に紐付く文書を削除します。
2. 前回のクローラ実行日以降に掲載期間が終了した掲示情報の文書を削除します。

#### 3.2.1.3 登録情報

掲示板クローラでは、掲示情報を以下の内容で登録します。

※全文検索時に検索対象となるフィールドはテキスト (text\_ngram または text\_morph)です。

Solr スキーマ定義に関する詳細は、「IM-ContentsSearch プログラミングガイド」 3.1 イントラマート標準 Solr スキーマ定義 をご参照ください。

項番	フィールド名	格納されるデータ	設定値	備考
1	id	文書を一意に識別する ID	isp_bulletin_ 掲示情報コード	
2	type	文書のタイプ	startpack/isp_bulletin/ カ テ ゴ リ (1..n)* / 種別	*種別に紐付く全ての カテゴリまたはコミュニ ティを登録
3	url	詳細画面 URL	startpack/bulletin/ntc_dtl_cmnt.jsp	掲示板内容表示画面
4	id_original	詳細画面を表示するために必要な パラメータ	・掲示板種別コード ・掲示情報コード ・ポータルフラグ ・コミュニティフラグ	
5	title	タイトル	タイトル	
6	text_ngram	N-gram 用テキストデータ	・タイトル ・内容 ・投稿者 【添付ファイルがある場合】 ・ファイル名 ・ドキュメント内テキスト※	N-gram を使用する場 合のみ設定  ※ファイルフォーマット がサポート対象の場合
7	text_morph	形態素解析用テキストデータ	上記に同じ	形態素解析を使用す る場合のみ設定
8	sids_allowed	閲覧可能権限	次章参照	
9	sids_denied	閲覧不可権限	設定なし	
10	record_date	登録日	文書の登録日時	
11	*_string	文書タイプ固有の文字列データ		掲示板固有のデータ
12		insert_user_name_string	投稿者	
13		insert_date_string	投稿日	
14		update_date_string	更新日	
15		full_class_name_string	種別名(階層付)	
16		cls_end_date_string	種別の終了日付	
17		info_end_date_string	掲載期間終了日	

### 3.2.1.4 アクセス権

掲示板	掲示情報に紐付く掲示板種別のアクセス権を基に、以下の権限が 付与されたグループ※1 が閲覧可能権限に設定されます。 ・管理、登録可、参照可
コミュニティ掲示板	コミュニティメンバを閲覧可能権限に設定します。※2

※1 会社・組織に役職による絞込みがある場合は、該当ユーザが登録されます。

※2 基準となる日: クローラ実行日

### 3.2.2 業務テンプレート表示情報

- タイトル
- 階層付種別名
- 投稿者

- 投稿日
- 要約(要約表示の指定がある場合)

<b>【掲示板】 勤怠締切のお知らせ</b>	お知らせ/事務関連情報
投稿者: 関根千香	投稿日: 2008/07/07
来月の <del>勤怠申請締切日</del> は来月の5日です。	

<掲示板テンプレートイメージ>

### 3.2.3 コンテンツ画面

回 交通費支給方法の変更について				最新情報 ? ヘルプ 閉じる	
投稿者: 松本 英樹 2008/07/07(月) 12:00 種別: お知らせ/事務関連情報					
<p>今月から交通費支給の方法が変わります。 詳細はメールにてご連絡いたします。</p>					
添付ファイル	ファイル名	サイズ	登録日時		
掲載期間	指定なし				

<掲示板コンテンツ画面イメージ>

## 3.3 電子会議室

### 3.3.1 クローラ

動的な文書種別情報の取得と電子会議室情報の Solr サーバへの登録・更新・削除を行います。  
※動的な文書種別情報とは、全文検索結果画面の種別の絞込みに使用されるカテゴリ・種別情報です。

#### 3.3.1.1 登録・更新対象

<<基本条件>>

- ・クローラ実行日時時点で有効な会議室種別またはコミュニティ会議室種別の電子会議室情報を対象とします。

#### ■ 再作成クローラ(再作成範囲指定なし)

基本条件を満たした範囲期間内の全ての電子会議室情報を対象とします。

範囲期間に関しては、3.1.1.3 再作成クローラが対象とする範囲 に準じます。(以下同様)

#### ■ 差分クローラ、再作成クローラ(再作成範囲指定あり)

基本条件を満たし、かつ以下の条件のいずれかを満たす電子会議室情報を対象とします。

1. 前回のクローラ実行日または再作成日以降に会議室種別が有効になった種別に紐付く更新日が初期設定日以降の電子会議室情報

2. 前回のクローラ実行日または再作成日以降にマスタ情報が更新されたカテゴリ・種別に紐付く更新日が初回設定日以降の電子会議室情報
3. 前回のクローラ実行日または再作成指定日以降に登録・更新された電子会議室情報
4. 前回のクローラ実行日または再作成指定日以降にコミュニティに参加したメンバがいるコミュニティの更新日が初回設定日以降の電子会議室情報

### 3.3.1.2 削除対象

- 再作成クローラ  
範囲期間内の電子会議室に紐付く全ての文書を削除します。
- 差分クローラ  
前回のクローラ実行日以降に有効期限切れになった会議室種別に紐付く文書を削除します。

### 3.3.1.3 登録情報

電子会議室クローラでは、会議室情報を以下の内容で登録します。

※全文検索時に検索対象となるフィールドはテキスト (text\_ngram または text\_morph)です。

項番	フィールド名	格納されるデータ	設定値	備考
1	id	文書を一意に識別する ID	isp_conference_ 掲示情報コード	
2	type	文書のタイプ	startpack/ isp_conference /カテゴリ (1..n)* /種別	*種別に紐付く全ての カテゴリまたはコミュニ ティを登録
3	url	詳細画面 URL	startpack/conference/mtr_dtl_cmnt.j ssp	電子会議室参照画面
4	id_original	詳細画面を表示するために必要な パラメータ	・会議室種別コード ・会議室情報コード ・ポータルフラグ ・コミュニティフラグ	
5	title	タイトル	タイトル	
6	text_ngram	N-gram 用テキストデータ	・タイトル ・内容 ・投稿者 【添付ファイルがある場合】 ・ファイル名 ・ドキュメント内テキスト※	N-gram を使用する場 合のみ設定  ※ファイルフォーマット がサポート対象の場合
7	text_morph	形態素解析用テキストデータ	上記に同じ	形態素解析を使用す る場合のみ設定
8	sids_allowed	閲覧可能権限	次章参照	
9	sids_denied	閲覧不可権限	設定なし	
10	record_date	登録日	文書の登録日時	
11	*_string	文書タイプ固有の文字列データ		電子会議室固有のデ ータ
12	insert_user_name_string		投稿者	
13	insert_date_string		投稿日	
14	update_date_string		更新日	
15	full_class_name_string		種別名(階層付)	
16	cls_end_date_string		種別の終了日付	

## 3.3.1.4 アクセス権

電子会議室	会議室情報に紐づく会議室種別のアクセス権を基に、以下の権限が付与されたグループ※1 が閲覧可能権限に設定されます。 ・管理、登録可、参照可
コミュニティ電子会議室	コミュニティメンバを閲覧可能権限に設定します。※2

※1 会社・組織に役職による絞込みがある場合は、該当ユーザが登録登録されます。

※2 基準となる日:クローラ実行日

## 3.3.2 業務テンプレート表示情報

- タイトル
- 階層付種別名
- 投稿者
- 投稿日
- 要約(要約表示の指定がある場合)

【電子会議室】販売ガイド 営業部/新製品関連  
 投稿者: 関根千香 投稿日: 2008/07/07  
 新製品の販売 マニュアルを作成しました。

<電子会議室テンプレートイメージ>

## 3.3.3 コンテンツ画面

電子会議室 参照画面

最新情報 ヘルプ 閉じる

交流会について

投稿者: 松本 英樹 1970/01/01(木) 09:00 種別: 社内関連/アンケート

次月に行われる交流会ですが、なにか要望等はございますか。

添付ファイル	ファイル名	サイズ	登録日時
コメント投稿			
タイトル	投稿者	投稿日時	
交流会について	松本 英樹		
Re:交流会について	武藤 美代子		
Re:Re:交流会について	松本 英樹		

<電子会議室コンテンツ画面イメージ>

## 3.4 ドキュメント管理

### 3.4.1 クローラ

動的な文書種別情報の取得とドキュメント管理ドキュメント情報の Solr サーバへの登録・更新・削除を行います。  
※動的な文書種別情報とは、全文検索結果画面の種別の絞込みに使用されるフォルダ情報です。

#### 3.4.1.1 登録・更新対象

<<基本条件>>

・クローラ実行日時時点で参照が可能なドキュメント管理ドキュメント情報またはドキュメント管理ドキュメント情報を対象とします。

ドキュメントの登録時にワークフローの承認が必要な場合、未承認のドキュメント情報は登録対象外です。

##### ■ 再作成クローラ(再作成範囲指定なし)

基本条件を満たし、範囲期間内の全てのドキュメント管理ドキュメント情報を対象とします。

範囲期間に関しては、3.1.1.3 再作成クローラが対象とする範囲 に準じます。(以下同様)

##### ■ 差分クローラ、再作成クローラ(再作成範囲指定あり)

基本条件を満たし、かつ以下の条件のいずれかを満たすドキュメント管理ドキュメント情報を対象とします。

1. 前回のクローラ実行日または再作成日以降にワークフローで登録承認済みになったドキュメント管理ドキュメント情報
2. 前回のクローラ実行日または再作成日以降にマスタ情報が更新されたフォルダに紐付く更新日が初回設定日以降のドキュメント管理ドキュメント情報
3. 前回のクローラ実行日または再作成日以降に登録・更新されたドキュメント管理ドキュメント情報
4. 前回のクローラ実行日または再作成日以降にコミュニティに参加したメンバがいるコミュニティの更新日が初回設定日以降のドキュメント管理ドキュメント情報
5. 関連ファイルが削除されたまたは前回のクローラ実行日または再作成日以降に関連ファイルが更新された更新日が初回設定日以降のドキュメント管理ドキュメント情報

#### 3.4.1.2 削除対象

##### ■ 再作成クローラ

範囲期間内のドキュメント管理に紐付く全ての文書を削除します。

##### ■ 差分クローラ

前回のクローラ実行日以降にワークフローで削除承認された文書を削除します。

#### 3.4.1.3 登録情報

ドキュメント管理クローラでは、ドキュメント管理ドキュメント情報を以下の内容で登録します。

※全文検索時に検索対象となるフィールドはテキスト (text\_ngram または text\_morph) です。

項番	フィールド名	格納されるデータ	設定値	備考
1	id	文書を一意に識別する ID	isp_document ドキュメントコード	
2	type	文書のタイプ	startpack/isp_document/ フォルダ (1..n)*	*親フォルダを含めた全てのフォルダまたはコミュニティを登録

3	url	詳細画面 URL	startpack/document/doc_dtl_doc.jsp	ドキュメント参照画面
4	id_original	詳細画面を表示するために必要なパラメータ	<ul style="list-style-type: none"> <li>・フォルダコード</li> <li>・ドキュメントコード</li> <li>・ポータルフラグ</li> <li>・コミュニティフラグ</li> </ul>	
5	title	タイトル	ドキュメント名	
6	text_ngram	N-gram 用テキストデータ	<ul style="list-style-type: none"> <li>・ドキュメント名</li> <li>・備考</li> <li>【添付ファイルがある場合】</li> <li>・ファイル名</li> <li>・ドキュメント内テキスト※</li> <li>【関連ドキュメントがある場合】</li> <li>・ドキュメント名</li> </ul>	N-gram を使用する 場合のみ設定  ※ファイルフォーマット がサポート対象の場合
7	text_morph	形態素解析用テキストデータ	上記に同じ	形態素解析を使用する 場合のみ設定
8	sids_allowed	閲覧可能権限	次章参照	
9	sids_denied	閲覧不可権限	次章参照	
10	record_date	登録日	文書の登録日時	
11	*_string	文書タイプ固有の文字列データ		ドキュメント管理固有の データ
12		insert_user_name_string	登録者	
13		insert_date_string	登録日	
14		update_date_string	更新日	
15		full_folder_name_string	フォルダ名 (階層付)	
17				
18				

#### 3.4.1.4 アクセス権

フォルダ	ドキュメント管理ドキュメント情報に紐づくフォルダのアクセス権を基に、以下の権限が付与されたグループ※1 が閲覧可能権限に設定されます。 ・管理、登録可、参照可
コミュニティフォルダ	コミュニティメンバを閲覧可能権限に設定します。※2

※1 会社・組織に役職による絞込みがある場合は、該当ユーザが登録されます。

※2 基準となる日: クローラ実行日

#### 3.4.2 業務テンプレート表示情報

- ドキュメント名
- 階層付フォルダ名
- 登録者
- 登録日
- 要約 (要約表示の指定がある場合)



【ドキュメント管理】 原価情報 営業社外秘/価格表  
 登録者: 安達 健太 登録日: 2008/07/07  
 設計20製造55結合テスト40 **テスト**結合

<ドキュメント管理テンプレートイメージ>

### 3.4.3 コンテンツ画面

最新情報 ヘルプ 閉じる			
フォルダ	総務		
ドキュメント名	交通費申請		
添付ファイル	ファイル名	サイズ	登録日時
	様式19交通費請求明細書(企画部対応).xls	18KB	2009/05/20(水) 16:38:58
関連ドキュメント	フォルダ名	ドキュメント名	
登録者	坂田 正行	登録日時	2009/05/20(水) 16:38:58
最終更新者	坂田 正行	最終更新日時	2009/05/20(水) 16:38:58
備考			
ソート番号	0		

<ドキュメント管理コンテンツ画面イメージ>

## 3.5 スケジュール

### 3.5.1 クローラ

Solr サーバへスケジュール情報の登録・更新・削除を行います。

#### 3.5.1.1 登録・更新対象

- 再作成クローラ(再作成範囲指定なし)  
 範囲期間内の全てのスケジュール情報を対象とします。  
 範囲期間に関しては、3.1.1.3 再作成クローラが対象とする範囲 に準じます。(以下同様)
- 差分クローラ、再作成クローラ(再作成範囲指定あり)  
 基本条件を満たし、かつ以下の条件のいずれかを満たすスケジュール情報を対象とします。
  1. 前回のクローラ実行日または再作成日以降に登録・更新されたスケジュール情報
  2. 前回のクローラ実行日または再作成日以降に個人アクセス権に登録・更新したユーザが参加する更新日が初回設定日以降のスケジュール情報
  3. 前回のクローラ実行日または再作成日以降にグループアクセス権に登録・更新した組織・グループに所属するユーザが参加する更新日が初回設定日以降のスケジュール情報

#### 3.5.1.2 削除対象

- 再作成クローラ  
 範囲期間内のスケジュールに紐付く全ての文書を削除します。
- 差分クローラ  
 なし

## 3.5.1.3 登録情報

スケジュールクローラでは、スケジュール情報を以下の内容で登録します。

※全文検索時に検索対象となるフィールドはテキスト (text\_ngram または text\_morph) です。

項番	フィールド名	格納されるデータ	設定値	備考
1	id	文書を一意に識別する ID	isp_schedule_予約コード	
2	type	文書のタイプ	startpack/isp_schedule/年/年月	年: isp_schedule_yyyy 年月: isp_schedule_yyyyMM
3	url	詳細画面 URL	startpack/contents_search/sdl_detail. jssp	
4	id_original	詳細画面を表示するために必要な パラメータ	・親画面 ID(CNTS) ・予約コード ・ログインユーザコード	
5	title	タイトル	タイトル	
6	text_ngram	N-gram 用テキストデータ	・タイトル ・開始日付 (yyyy/MM/dd) ・終了日付 (yyyy/MM/dd) ・参加者名 ・説明 ・場所 ・来訪者欄の会社名、代表者名 【添付ファイルがある場合】 ・ファイル名 ・ドキュメント内テキスト※	N-gram を使用する場 合のみ設定  ※ファイルフォーマット がサポート対象の場合
7	text_morph	形態素解析用テキストデータ	上記に同じ	形態素解析を使用す る場合のみ設定
8	sids_allowed	閲覧可能権限	次章参照	
9	sids_denied	閲覧不可権限	次章参照	
10	record_date	登録日	文書の登録日時	
11	*_string	文書タイプ固有の文字列データ		スケジュール固有のデ ータ
12		start_date_string	開始日時	
13		end_date_string	終了日時	
14		place_string	場所	
15		title_string	タイトル	
16		participants_string_mlt	参加者(マルチバリュー)	
17		all_day_flg_string	終日フラグ	
18		insert_user_name_string	登録者	
18		update_date_string	更新日	

## 3.5.1.4 アクセス権

スタートパックのスケジュールのアクセス権は非常に複雑な構造となっており、そのまま IM-ContentsSearch のアクセス権に適用することができない為、全文検索におけるスケジュールの閲覧可能ユーザは以下に制限します。

- ・ スケジュールの参加者
- ・ スケジュールの登録者
- ・ スケジュールの参加者に対して秘書権を持つユーザまたはグループ※

※ 会社・組織に役職による絞込みがある場合は、該当ユーザが登録。

参加者内に、以下の条件に当てはまるユーザが存在しない場合は、閲覧可能権限に該当ユーザまたはグループが設定されます。

- ・ 個人アクセス権の設定で、全体設定を「秘書権」にしている
- ・ 参加者が所属するグループアクセス権の設定で、全体設定を「秘書権」にしている

条件に当てはまるユーザが存在する場合は、閲覧可能権限には「全ユーザ」、閲覧不可権限には閲覧可能権限を持たないユーザが設定されます。

アクセス権を設定する際に組織・パブリックグループの所属情報を参照する場合は、クローラ実行日時点の所属情報を参照します。

### 3.5.2 業務テンプレート表示情報

- タイトル
- 日時(開始～終了)
- 場所
- 参加者
- 登録日
- 要約(要約表示の指定がある場合)

#### 【スケジュール】七幡製作所との会議

日時: 2009/06/01 10:00 - 2009/06/01 12:00      場所: 会議室

参加者: 加藤 隆, 澤島 誠二, 福田 雄太

登録者: 根本 真理

**七幡製作所**の高橋部長が来社されます。

<スケジュールテンプレートイメージ>

### 3.5.3 コンテンツ画面

スケジュール 参照
 
 ヘルプ
 閉じる

基本情報			
タイトル	七幡製作所との会議		
期間	2009/06/01(月) 10:00 から 2009/06/01(月) 12:00		
参加者	福田 雄太, 加藤 隆, 澤島 誠二		
施設・備品	会議室		
説明	七幡製作所の高橋部長が来社されます。		
場所	会議室		
外出			

添付ファイル	ファイル名	サイズ	登録日時
<div> <div>詳細情報</div> <div> <div> <div>登録者</div> <div>IM製造 / 営業部 / マーケティングチーム 根本 真理</div> </div> <div> <div>更新者</div> <div>IM製造 / 営業部 / マーケティングチーム 根本 真理</div> </div> </div> <div> <div> <div>E-mail</div> <div>nemoto@localhost</div> </div> <div> <div>E-mail</div> <div>nemoto@localhost</div> </div> </div> <div> <div>重要度</div> <div>中</div> </div> <div> <div>登録状態</div> <div>仮</div> </div> <div> <div>来訪者</div> <div></div> </div> </div> <div> <div>非表示</div> </div>			

添付ファイル

ファイル名

サイズ

登録日時

<スケジュールコンテンツ画面イメージ>

## 3.6 ブログ記事検索

### 3.6.1 クローラ

動的な文書種別の取得とブログ記事情報の Solr サーバへの登録・更新・削除を行います。

※動的な文書種別情報とは、全文検索結果画面の種別の絞込みに使用されるカテゴリ・種別情報です。

#### 3.6.1.1 登録・更新対象

##### ■ 再作成クローラ(再作成範囲指定なし)

範囲期間内の全てのブログ記事情報を対象とします。

範囲期間に関しては、3.1.1.3 再作成クローラが対象とする範囲 に準じます。(以下同様)

##### ■ 差分クローラ、再作成クローラ(再作成範囲指定あり)

以下の条件のいずれかを満たすブログ記事情報を対象とします。

1. 前回のクローラ実行日または再作成日以降に登録・更新されたブログ記事情報
2. 前回のクローラ実行日または再作成日以降にマスタ情報が更新されたカテゴリ、種別に紐付く更新日が初回設定日以降のブログ記事情報

#### 3.6.1.2 削除対象

##### ■ 再作成クローラ

範囲期間内のブログ記事に紐付く全ての文書を削除します。

##### ■ 差分クローラ

なし

#### 3.6.1.3 登録情報

ブログ記事クローラでは、ブログ記事情報を以下の内容で登録します。

※全文検索時に検索対象となるフィールドはテキスト (text\_ngram または text\_morph)です。

項番	フィールド名	格納されるデータ	設定値	備考
1	id	文書を一意に識別する ID	isp_blog_記事コード	
2	type	文書のタイプ	startpack/isp_blog/カテゴリ/種別	
3	url	詳細画面 URL	startpack/portal/ptl_blg_atcl_dtl.jssp	記事詳細画面
4	id_original	詳細画面を表示するために必要なパラメータ	記事コード	
5	title	タイトル	タイトル	
6	text_ngram	N-gram 用テキストデータ	・タイトル ・カテゴリ名 ・種別名 ・内容 【添付ファイルがある場合】 ・ファイル名 ・ドキュメント内テキスト※	N-gram を使用する 場合のみ設定  ※ファイルフォーマット がサポート対象の場合
7	text_morph	形態素解析用テキストデータ	上記に同じ	形態素解析を使用する 場合のみ設定
8	sids_allowed	閲覧可能権限	ALLユーザ	
9	sids_denied	閲覧不可権限	なし	

10	record_date	登録日	文書の登録日時	
11	*_string	文書タイプ固有の文字列データ		ブログ記事固有のデータ
12		category_name_string	カテゴリ名	
13		kind_name_string	種別名	
14		insert_date_string	登録日時	
15		insert_user_name_string	登録者	
16		update_date_string	更新日	
17				

### 3.6.2 業務テンプレート表示情報

- タイトル
- カテゴリ名/種別名
- 登録者
- 登録日
- 要約(要約表示の指定がある場合)

**【ブログ記事】 環境問題について**      生活・健康/家庭用品  
 登録者: 安達 健太      登録日: 2008/07/07  
 最近 **環境** に対する関心がますます高まって

<ブログ記事テンプレートイメージ>

### 3.6.3 コンテンツ画面

記事詳細 最新情報 閉じる

▶ 環境問題について 生活・健康/家庭用品  
 登録日 2008/07/07(月) 12:00

最近 **環境** に対する関心がますます高まっていますが、映画でも環境破壊を危惧する内容のものが多くなったように感じます。そこで俄然然になるのが、近日公開となるあの映画です。ロードショー当日に映画館に足を運んでみるつもりです。映画を見て環境問題に対し関心を持つようになった私、少しでも力になろうと思い、近頃はエコバッグを携帯しています。自分ができることから、無理なく始めてみるのも悪くないですね。

<記事ID: adachi\_blog\_1>  
アクセス履歴 (1)

▶ 関連記事

▶ コメント

コメント投稿 投稿

<ブログ記事コンテンツ画面イメージ>

## 3.7 営業日報

### 3.7.1 クローラ

動的な文書種別情報の取得と営業日報情報の Solr サーバへの登録・更新・削除を行います。

※動的な文書種別情報とは、全文検索結果画面の種別の絞込みに使用される情報です。

#### 3.7.1.1 登録・更新対象

- 再作成クローラ(再作成範囲指定なし)  
範囲期間内の全ての営業日報情報を対象とします。  
範囲期間に関しては、3.1.1.3 再作成クローラが対象とする範囲 に準じます。(以下同様)
- 差分クローラ、再作成クローラ(再作成範囲指定あり)  
以下の条件のいずれかを満たす営業日報情報を対象とします。
  1. 前回のクローラ実行日または再作成日以降に登録・更新された営業日報情報。
  2. 前回のクローラ実行日または再作成日以降にマスタ情報が更新された顧客情報に紐付く更新日が初回設定日以降の営業日報情報

#### 3.7.1.2 削除対象

- 再作成クローラ  
範囲期間内の全ての営業日報情報を削除します。
- 差分クローラ  
なし

#### 3.7.1.3 登録情報

営業日報クローラでは、営業日報情報を以下の内容で登録します。

※全文検索時に検索対象となるフィールドはテキスト (text\_ngram または text\_morph)です。

項番	フィールド名	格納されるデータ	設定値	備考
1	id	文書を一意に識別する ID	isp_daily_ 日報 ID	
2	type	文書のタイプ	startpack/ isp_daily /登録者/実施年月	登録者: isp_daily_%user_cd% 実施年月: isp_daily_yyyyMM
3	url	詳細画面 URL	startpack/report/bsn_sdl_ref.jsp	営業日報参照画面
4	id_original	詳細画面を表示するために必要なパラメータ	日報 ID ポータルフラグ	
5	title	タイトル	タイトル	
6	text_ngram	N-gram 用テキストデータ	・タイトル ・レポート内容 ・実施日 (yyyy/MM/dd) 【添付ファイルがある場合】 ・ファイル名 ・ドキュメント内テキスト※	N-gram を使用する 場合のみ設定 ※ファイルフォーマット がサポート対象の場合
7	text_morph	形態素解析用テキストデータ	上記に同じ	形態素解析を使用す

				る場合のみ設定
8	sids_allowed	閲覧可能権限	次章参照	
9	sids_denied	閲覧不可権限	なし	
10	record_date	登録日	文書の登録日時	
11	*_string	文書タイプ固有の文字列データ		営業日報固有のデータ
12		start_date_string	開始日時	
13		end_date_string	終了日時	
14		customer_name_string	顧客名	
15		importance_string	重要度	
16		insert_date_string	登録日時	
17		insert_user_name_string	登録者	
18		update_date_string	更新日	

### 3.7.1.4 アクセス権

日報作成者の所属組織と配信先ユーザを閲覧可能権限として設定します。

- ・ 日報作成者の所属組織  
クローラ実行日時点の日報作成者が現在・過去・未来に所属する組織が全て登録されます。
- ・ 配信先設定先ユーザ  
営業日報登録時に設定された配信先ユーザ・グループに所属するユーザが登録されます。

## 3.7.2 業務テンプレート表示情報

- タイトル
- 日報記入者
- 実施日時(開始～終了)
- 顧客名(SFA)
- 要約(要約表示の指定がある場合)

### 【営業日報】 東鉄製作所 製品説明

日報記入者: 澤島 誠二      実施日時: 2009/05/01(金) 13:00 - 15:00

顧客名(SFA): 東鉄製作所.

東鉄製作所 製品説明。

<営業日報業務テンプレートイメージ>

## 3.7.3 コンテンツ画面

営業日報参照				 ヘルプ  閉じる	
日報記入者	澤島 誠二				
更新日時	2009/06/09(火) 15:11		最終配信日	2009/07/07(火)	
タイトル	東鉄製作所 製品説明				
顧客名(SFA)				顧客担当	
実施日時	2009/05/01(金) 13:15 - 15:15				
次回訪問日					
重要度	WARM				
レポート内容	東鉄製作所 製品説明。				
添付ファイル	ファイル名	サイズ	登録日時		
配信設定	配信あり				
コメント					

&lt;営業日報コンテンツ画面イメージ&gt;



## 4 ISP クローラの拡張

ISP クローラでは、ISP クローラ用のリスナーを作成して追加することで、登録情報として独自の項目を追加することが可能です。

この章では、独自の項目を登録情報に追加するためのリスナーの作成方法について説明します。

### 4.1 リスナーの呼び出し

ISP クローラは情報単位で Solr サーバに文書を登録していきますが、登録処理の直前に設定ファイルに定義されたリスナーを呼び出し、リスナーが存在した場合はリスナーの項目追加メソッドを実行します。

リスナーが ISP クローラに呼び出されるシーケンスを以下に示します。

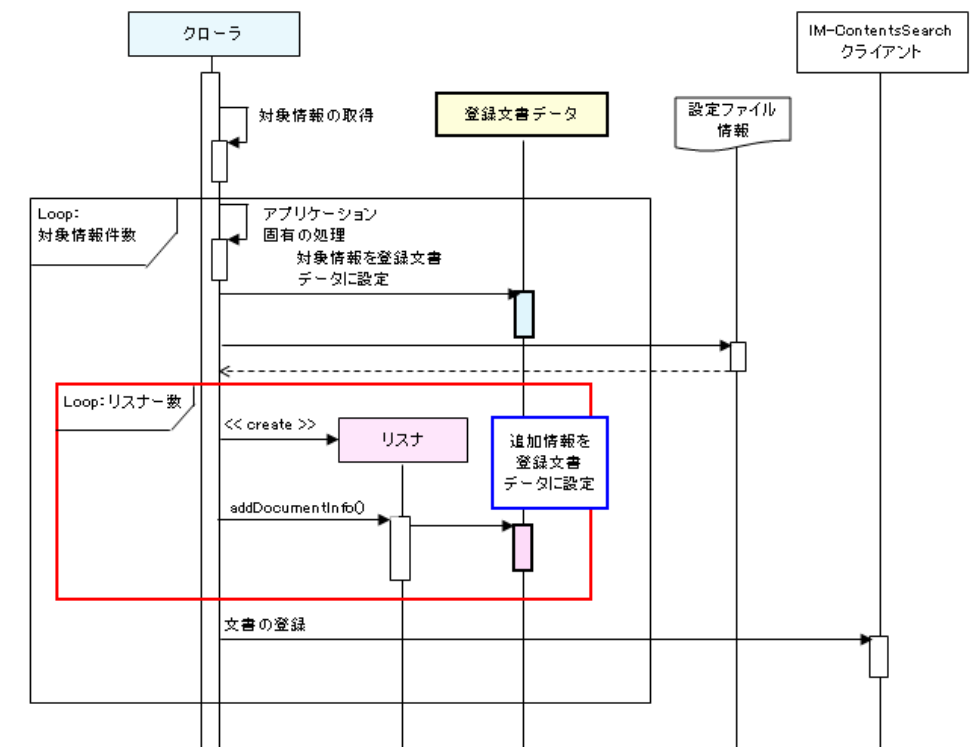


図 4.1-1 リスナー呼び出しシーケンス

### 4.2 リスナの作成と設定

各機能の拡張用インタフェースを実装したリスナーを作成します。

作成したリスナクラスは、以下のファイルの<listener>タグ内に該当機能のタグを追加し listener-class に指定します。

```
Server Manager/conf/ system-install-isp-sfa-comm.xml
```

(掲示板クローラ用リスナ設定例)

```
<listener>
.....
<bulletincrawler-add-listener>
  <listener-class> × × × × × </listener-class>
</bulletincrawler-add-listener>
</listener>
```

機能	タグ	リスナが実装するインタフェース
		jp.co.intra_mart.product.startpack.api.contents_search.listener
掲示板	bulletincrawler-add-listener	BulletinCrawlingAddListener
電子会議室	conferencecrawler-add-listener	ConferenceCrawlingAddListener
ドキュメント管理	documentcrawler-add-listener	DocumentCrawlingAddListener
スケジュール	schedulecrawler-add-listener	ScheduleCrawlingAddListener
ブログ記事	blogcrawler-add-listener	BlogCrawlingAddListener
営業日報	dailycrawler-add-listener	DailyCrawlingAddListener

#### 4.2.1 処理の記述

登録情報に独自の項目を追加するための実処理は **addDocumentInfo** メソッド内に記述します。

**addDocumentInfo** メソッドの引数には、文書登録用オブジェクト、キー情報が渡されます。

リスナーでは以下の処理を行います。

- ① 引数のキー情報を元に、追加データを取得する。
- ② 引数の文書登録用オブジェクトにアプリケーションデータを追加する。
  - ・ 検索対象となるテキストフィールドにデータを登録したい場合  
IntramartSolrInputDocument #inputDocumentAddText メソッドを使用
  - ・ 表示目的等の格納用データを登録したい場合  
IntramartSolrInputDocument # addField メソッドを使用してダイナミックフィールドにデータを追加します。

※いずれのフィールド値にも null 値を設定することは出来ません。

以下に、掲示情報に追加情報を登録するサンプルを示します。

```
public class BulletinCrawlingSampleListener implements BulletinCrawlingAddListener {
略.....

/**
 * 登録情報に追加データを設定します。
 *
 * @param inputDoc
 *         入力文書
 * @param bulletinClsCd
 *         掲示板種別コード
 * @param bulletinInfoCd
 *         掲示情報コード
 * @throws SolrCrawlerException
 *         リスナーの処理で例外が発生した場合
 */
public void addDocumentInfo(IntramartSolrInputDocument inputDoc,
    String bulletinClsCd, String bulletinInfoCd)
    throws SolrCrawlerException {

    try {
        // 引数のキー情報を元にデータを取得する
        .....

        // 1. テキストフィールドにデータを追加
        inputDoc.addText("追加データ TEXT1");
        inputDoc.addText("追加データ TEXT2");

        // 2. ダイナミックフィールドにデータを追加
        inputDoc.addField("add_data1_string","追加データ_ダイナミックフィールド1");
        inputDoc.addField("add_data2_string","追加データ_ダイナミックフィールド2");
    } catch (Exception e) {
        throw new SolrCrawlerException(e.getMessage(), e);
    }
}
```

IM-ContentsSearch for ISP Ver.7.0  
ISP クローラ仕様書

2009/09/30 初版

Copyright 2000-2009 株式会社 NTT データ イントラマート  
All rights Reserved.

TEL: 03-5549-2821

FAX: 03-5549-2816

E-MAIL: [info@intra-mart.jp](mailto:info@intra-mart.jp)

URL: <http://www.intra-mart.jp/>