

自治体行政情報流通連携基盤システムの 構築・動作確認方法

1.1.1. 本ドキュメントに示す自治体行政情報流通連携基盤システムの構成

本ドキュメントに示す自治体行政情報流通連携基盤システムは、以下の機能を持つ。

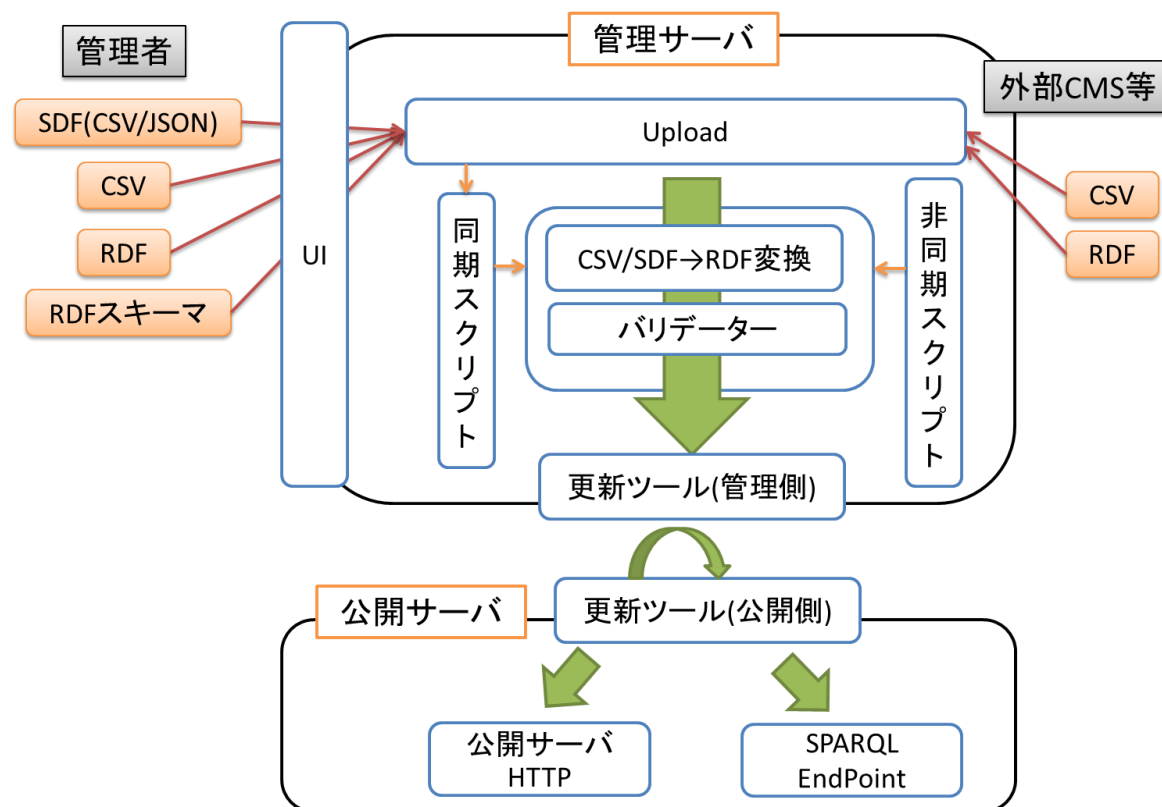
- ・ 管理者(自治体側の管理者)が RDF,CSV,SDF,RDF スキーマを登録する。
- ・ CSV,SDF に関しては、サーバ上で RDF に変換/バリデート/値の型付を行う。
- ・ 変換された RDF,RDF スキーマは公開サーバ（複数）が管理サーバから取得する
- ・ 管理サーバは取得した RDF を SPARQL EndPoint に格納し、利用者（市民）が SPARQL で検索可能にする。
- ・ 管理サーバは取得した RDF,RDF スキーマを HTTP で利用者が閲覧可能な状態にする。

なお、

- ・ 管理サーバ：自治体の担当職員のみが使用するデータ登録用サーバ
- ・ 公開サーバ：市民が使用するデータ公開用サーバ

である。

また、本ドキュメントでは、以下のシステム構成を想定している。



1.1.2. 構築手順

以下にシステム構築手順を示す。

なお、後述するサーバ構築に記載しているミドルウェア等以外に、別途以下のリソースが必要となる。

・Java ライブラリ(jar ファイル)

- ・ GSON (gson-*.jar)(<https://code.google.com/p/google-gson/>)
- ・ Apache Jena と Apache Jena が要求するライブラリ (<https://jena.apache.org/> から Apache Jena(apache-jena-*.tar.gz)をダウンロードし、その中の lib 以下の jar ファイル一式)

※上記 jar ファイルは変換ツールの Java ソースコードからのコンパイルにも必要である。

・RDF スキーマファイル

各データが参照するボキャブラリの RDF スキーマファイルを別途取得し、適切なファイル名、拡張子を付与の上、管理サーバ上に配置する必要がある。なお、取得した RDF スキーマファイルが文法上正しくない場合データの変換に失敗するため、必要に応じて取得した RDF スキーマファイルを修正する。

なお、以下の名前空間のボキャブラリはシステム側で参照しているため、作成データに関わらず RDF スキーマファイルが必要になる。

dcterms: <http://purl.org/dc/terms/>
dc: <http://purl.org/dc/elements/1.1/>
rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
rdfs: <http://www.w3.org/2000/01/rdf-schema#>

1.1.2.1.構築に関しての留意点

- ・ OS、ネットワーク設定、セキュリティ設定については本基盤システム固有の箇所以外の説明は省略する。導入環境のポリシーに従うこと。また、クラウド環境の場合は OS 展開済みのイメージを使用する場合もある
- ・ ミドルウェアについては、導入パッケージによりディレクトリ構成やオプション等が異なることがあるため、要点についてのみ記載する。
- ・ サーバ構成に記載しているミドルウェアは一例であり、別のプロダクトに変更可能な場合もあるが、本ドキュメントでは省略する。

1.1.2.2.公開サーバの構築

公開サーバの構築の流れの概略を以下に示す。

- ・ OS のインストール
- ・ ミドルウェアの導入設定
- ・ Jena Fuseki のインストール(公開サーバのみ)
- ・ 変換ツール導入
- ・ 個別設定

1.1.2.2.1. サーバ構築

以下の設定でサーバを構築する。

表 3-1-1-1-1 サーバ構成(公開サーバ)

項番	要素	内容	備考
1	OS	64bit Linux	bash,wget が動作する必要がある。 apache で proxy を使用するため、SELinux の設定を変更もしくは無効化する必要がある。
2	開放ポート	80(HTTP)	80 以外のポートは外部に対して閉じるが、必要に応じて作業用の SSH のポートを開放する。 3030 に関しては、B02 公開サーバデータ更新が内部で localhost でのアクセスで使用するため、設定に留意する。
3	HTTP サーバ	apache	mod_proxy,mod_rewrite が必要 自動起動は OFF にする
4	Java	Oracle Java7	
5	Ruby	Ruby 1.8 以上	ruby-devel irb も必要
6	ユーザ	loduser を作成	loduser が sudo 可能な設定にする。

以下は、ユーザ : loduser で作業する。

以下 sudo する必要のある箇所があるが、説明は省略する。

1.1.2.2.2. Jena Fuseki のインストール

Jena のサイト(<http://jena.apache.org/>) から、jena-fuseki のバイナリを取得、解凍し、解凍してできるディレクトリを fuseki にリネームして、/usr/lib/ の下にコピーする。

続いて、以下のファイルのパーミッションを変更する。

表 3-1-1-1-2 Jena Fuseki 関連ディレクトリ(公開サーバ)

項番	パス	パーミッション	内容
1	/usr/lib/fuseki/s-*	755	ツール群
2	/usr/lib/fuseki/fuseki	755	本体
3	/usr/lib/fuseki/fuseki-server	755	本体

サーバ起動時にデータの更新、fuseki の起動、apache の起動をするための設定を追加する。
/etc/rc.local の末尾に以下を追加する。

```
/home/loduser/jrrktool/runfuseki.sh
sleep 60
/home/loduser/jrrktool/updatetdb.sh
```

```
/etc/init.d/httpd start
```

1.1.2.2.3. 変換ツール導入

以下のディレクトリを作成する。

表 3-1-1-1-3-1 変換ツール 関連ディレクトリ(公開サーバ)

項番	パス	パーミッション	内容
1	/home/loduser/work	644	作業用
2	/home/loduser/jrrktool/	777	データ登録ツール格納先
3	/home/loduser/jrrktool/updatefile	777	データ登録ツール作業ディレクトリ
4	/var/log/jrrk/	777	データ登録ツールログ出力先
5	/var/log/fuseki	777	標準 API ログ出力先
6	/usr/lib/fuseki/DB2	777	データベースのデータ格納先

B02 公開サーバデータ更新モジュール、設定ファイルを導入する。

表 3-1-1-1-3-2 スクリプト等 (公開サーバ) (script/outsvr/)

項番	パス	パーミッション	内容
1	/home/loduser/jrrktool/updatetdb.sh	755	公開サーバデータ更新モジュール
2	/home/loduser/jrrktool/runfuseki.sh	755	fuseki の起動
3	/home/loduser/jrrktool/fusekiloop.sh	755	fuseki の起動、自動再起動
4	/etc/httpd/conf.d/fuseki.conf ※パスは環境により異なる場合がある	644	Apache の設定 (fuseki への proxy) ※Sparql Endpoint を構成するソフトウェアを変更する場合は、変更後のソフトウェアの設定に従う
5	/usr/lib/fuseki/log4j.properties	644	fuseki ロガー設定
6	/usr/lib/fuseki/jrrk-tdb.ttl	644	fuseki 設定

バッチ処理を cron に設定する。

表 3-1-1-1-3-3 cron 設定 (公開サーバ)

項番	設定内容	変更内容
1	cron (crontab -e で設定)	MAILTO="" * * * * * /home/loduser/jrrktool/updatetdb.sh

1.1.2.2.4. 個別設定

以下のファイルを環境に合わせて修正する。

表 3-1-1-1-3-4 環境依存ファイル (公開サーバ)

項番	パス	変更内容
1	/home/loduser/jrrktool/updatetdb.sh	管理サーバの URL(\$KANRISVR) 公開サーバからアクセス可能な URL を指定する。

なお、環境によっては、上記以外に以下のファイルの修正が必要な場合がある。

表 3-1-1-1-3-5 環境依存設定一覧（公開サーバ）

項番	設定内容	対象ファイル
1	公開サーバホスト名	/etc/httpd/conf/httpd.conf ただし、前段に負荷分散装置等を挟む場合は、そちらの設定に従うこと。
2	管理サーバホスト名	/home/loduser/jrrktool/updatetdb.sh
3	管理サーバへの取得 間隔	crontab -e で /home/loduser/jrrktool/updatetdb.sh の項目を編集する。
4	公開するデータの基 底 URI	管理サーバ側で指定

設定終了後、サーバを再起動する。

1.1.2.3. 管理サーバの構築

管理サーバの構築の流れの概略を以下に示す。

- ・ OS のインストール（本書では省略。クラウド環境の場合は OS 展開済みのイメージを使用する場合もある）
- ・ ミドルウェアの導入設定（本書では省略）
- ・ 変換ツール導入
- ・ 個別設定

1.1.2.3.1. サーバ構築

以下の設定でサーバを構築する。

表 3-1-1-2-1 サーバ構成（管理サーバ）

項番	要素	内容	備考
1	OS	64bit の Linux	bash が動作する必要がある。 PHP から Shell を呼び出している箇所があるので、SELinux の設定を無効化する必要がある。
2	開放ポート	80(HTTP) 443(HTTPS)	80,443 以外のポートは外部に対して閉じるが、必要に応じて作業用の SSH 等のポートを開放する。
3	HTTP サーバ	apache	mod_headers, mod_ssl, mod_rewrite が必要 自動起動は ON にする
4	Java	Oracle Java7	
5	PHP		php-mbstring も必要 導入後、以下を php.ini に設定 <i>expose_php=Off</i> <i>post_max_size = 20M</i> <i>upload_max_filesize = 20M</i> <i>default_charset = "UTF-8"</i> また、 <i>error_reporting</i> も、

			<i>error_reporting = E_ERROR</i> に変更するが、導入直後は障害切り分けのために変更しない。
6	ユーザ	adminuser を作成	adminuser が sudo 出来るようにする。

以下は、ユーザ：adminuser で作業する。

以下 sudo する必要のある箇所があるが、説明は省略する。

1.1.2.3.2. 変換ツール導入

以下のディレクトリを作成する。

表 3-1-1-2-2-1 変換ツール関連ディレクトリ（管理サーバ）

項番	パス	パーミッション	内容
1	/home/adminuser/work	644	作業用
2	/home/ adminuser/jrrktool/	777	データ登録ツール格納先
3	/var/log/jrrk/	777	データ登録ツールログ出力先
4	/home/ adminuser /jrrktool/jar/	755	データ登録ツールライブラリ

変換ツール、設定ファイルを導入する。

表 3-1-1-2-2-2 変換ツール導入（管理サーバスクリプト）(script/kanrisvr)

項番	パス	パーミッション	内容
1	/home/adminuser/jrrktool/runconvert.sh	755	データ変換（バッチ）
2	/home/adminuser/jrrktool/ runconvertsync.sh	755	データ変換（オンライン）
3	/etc/httpd/conf.d/manage.conf ※パスは環境により異なる場合がある	644	Apache の設定 (レスポンスヘッダ、ssl, mime 設定) 認証用のパスワードファ イル等は別途作成する。

表 3-1-1-2-2-3 変換ツール導入（管理サーバ）(HTML,PHP,データ登録ツールライブラリ)

項番	パス	パーミッショ ン	内容
1	/var/www/html/*	ファイル：644 ディレクトリ ：777	データ登録ツール HTML,PHP
2	/home/adminuser/jrrktool/jar/convtool.jar	644	変換ツール
3	/home/adminuser/jrrktool/jar/*.jar	644	変換ツールが使用する jar ファイル (GSON,Apache Jena)
4	/var/www/html/lodmanage/commonschema/ 以下	644	ダブリンコア等の共通の RDF スキーマファイル。

バッチ処理を cron に設定する。

表 3-1-1-2-2-4 cron 設定（管理サーバ）

項番	設定内容	変更内容
1	cron (crontab -e で設定)	MAILTO="" * * * * */home/adminuser/jrrktool/runconvert.sh

1.1.2.3.3. 個別設定

以下のファイルを環境に合わせて修正する。

表 3-1-1-2-2-4 環境依存ファイル（管理サーバ）

項番	パス	変更内容
1	/var/www/html/lodmanage/searchresult.php	管理サーバからの公開サーバの Sparql Endpoint の URL (\$endpoint) 生成する RDF の基底 URI※1 (\$ownns)
2	/home/adminuser/jrrktool/runconvert.sh	生成する RDF の基底 URI※1 (Java を呼んでいる箇所の後半の URL)
3	/home/adminuser/jrrktool/runconvertsync.sh	生成する RDF の基底 URI※1 (Java を呼んでいる箇所の後半の URL)
4	/var/www/html/lodmanage/ckanupload.php	外部からの公開サーバの Sparql Endpoint の URL (\$ckanSparqlEndpoint) 外部からの公開サーバの RDF 公開 URL (\$openRdfUrlBase) 生成する RDF の基底 URI※1 (\$ownns)
5	/var/www/html/cmsupload.php	\$allowipar (外部システム (CMS 等) サーバの IP)もしくは、 \$allowhostar (外部システム (CMS 等) サーバのホスト名)

※1 生成する RDF の基底 URI は、データ設計時に定義するが、通常は導入自治体のドメインで、オープンデータ用に定義された URL になる。

なお、環境によっては、上記以外に以下のファイルの修正が必要な場合がある。

表 3-1-1-2-3-5 環境依存設定一覧（管理サーバ）

項番	設定内容	対象ファイル
1	公開サーバホスト名	/var/www/html/lodmanage/ckanupload.php (管理サーバから解決出来る URL で指定) /var/www/html/lodmanage/searchresult.php (外部から解決出来る URL で指定)
2	管理サーバホスト名	/etc/httpd/conf/httpd.conf
3	バッチ変換スクリプトの実行間隔	crontab -e で /home/adminuser/jrrktool/runconvert.sh の項目を編集する。

4	公開するデータの基底 URI	/var/www/html/lodmanage/searchresult.php /var/www/html/lodmanage/ckanupload.php /home/adminuser/jrrktool/runconvert.sh /home/adminuser/jrrktool/runconvertsync.sh
5	アップロードの上限サ イズ	/etc/php.ini の upload_max_filesize と post_max_size
6	外部システム（CMS 等）のアップロード元 の IP もしくはドメイ ン	/var/www/html/cmsupload.php
7	公開サーバの追加・削 除時の設定	不要（管理サーバは公開サーバの情報を持たない。）

設定終了後、サーバを再起動する。

1.1.3. 自治体行政情報流通連携基盤システム構築後の動作確認手順

1.1.3.1. 公開サーバの動作確認手順

1.1.3.1.1. HTTP サーバ/Sparql Endpoint の動作確認

管理サーバにまだデータファイルをアップロードしない時点で、以下の URL にアクセスする。
なお以下の URL は SPARQL Based Command にアクセスして、公開サーバのデータベース上の
RDF の Triple を 10 件表示するものである。

*http://[公開サーバの公開側のホスト名]/data/sparql?
query=SELECT+*+WHERE+%7B+graph+%3Fg%7B+%3Fs+%3Fp+%3Fo+.+%7D%7DLIMIT
+10&output=text*
(実際には一行)

以下のように表示されれば正常に動作している。

```
-----
| s | p | o | g |
=====
-----
```

なお、管理サーバにすでにデータファイルを登録した場合、データファイルの中身が反映され、
以下のような表示になる場合があるが、この場合も正常に動作している。(出力の右側と下側を省
略)

```
-----
| s | p | o | g |
=====
| <http://www.example.org/r3> | <http://www.example.org/c15> | "18" | <http://localh
```


1.1.3.2. 管理サーバの動作確認手順

公開サーバ動作確認後、

[https://\[管理サーバホスト名\]/loduser/lodmanage/setupupdate.php](https://[管理サーバホスト名]/loduser/lodmanage/setupupdate.php)

にブラウザでアクセスし、apache の Basic 認証で設定したユーザ名、パスワードでログインする。

その後、RDF スキーマ登録画面でサンプルの RDFS ファイル(testschema.rdf)を登録する。

さらに、サンプルの CSV ファイル(test.txt)を、データ登録で同期処理にチェックを入れて登録する。

このとき、登録に成功すれば管理サーバ側は正常に動作している。

1.1.3.3. 公開サーバ・管理サーバの結合確認手順

管理サーバの動作確認手順終了後、以下の URL を開く。

http://[公開サーバの公開側のホスト名]/

```
api/v1/sparql?query=SELECT+%3Ftitle+%0D%0AWHERE+%7B%0D%0AGRAPH+%3Fg%0D%0A%7B%0D%0A%3Chttp%3A%2F%2Fexample.org%2Fem0%3E+%3Chttp%3A%2F%2Fpurl.org%2Fdc%2Felements%2F1.1%2Ftitle%3E+%3Ftitle%0D%0A%7D%7D%0D%0A&output=text
```

※上記は以下のクエリを URL エンコードしたものと同等。

```
SELECT  ?title
WHERE {
  GRAPH ?g
  {
    <http://example.org/em0> <http://purl.org/dc/elements/1.1/title> ?title
  }
}
```

以下が表示されれば問題ない。

```
| title |
=====
| "jrrk" |
```

管理サーバで変換されたデータが、公開サーバ上のデータベースに格納され、そのデータが検索されたので、管理サーバと公開サーバで連係が行われている。

また、公開サーバ上で、管理サーバで変換した RDF ファイル ([http://\[公開サーバの公開側のホスト名\]/rdf/test.rdf](http://[公開サーバの公開側のホスト名]/rdf/test.rdf))、管理サーバに登録した RDFS ファイル([http://\[公開サーバの公開側のホスト名\]/schema/testschema.rdf](http://[公開サーバの公開側のホスト名]/schema/testschema.rdf))が取得できれば連携が行われている。

1.1.4. 自治体行政情報流通連携基盤システム運用における制限・留意事項

- RDF,RDFS,CSV,SDF のファイル名は、半角英数、アンダーバー(_)、ハイフン(-)のみとする。全角文字、スペースは使用できない。
- SDF に関して、SDF の仕様を暫定的にサポートし、かつボキャブラリの URI を定義するため独自拡張した SDF のみの対応とする。
- CSV に関しては、NameSpace,ボキャブラリの URI 等を付与した CSV を対象とする。作成

方法については データ CSV 化手法.pdf を参照のこと。なお、この CSV の仕様は、オープンデータ流通推進コンソーシアム 技術委員会 第 2 回委員会 (2013/1/9) 資料 2-4 技術委員会の論点 (<http://www.opendata.gr.jp/committee/docs/siryo2-4.pdf>) の P73 に記載されている「オープンデータ化のための CSV 形式データ規格案」を元にしたものである。

- ・バリデート対象外 URI ファイルの作成・メンテナンスが必要である。(管理サーバ：
/var/www/html/lodmanage/ignorelist/ignorelist.txt)
- ・バリデート対象外 URI ファイルは、バックアップファイルの同一ディレクトリへの格納はできない。(ディレクトリ中のファイルをすべて読むため)
- ・バリデート対象外 URI ファイル中でワイルドカードを使用する場合はドメインベースになるため、たとえば <http://www.w3.org/2000/01/rdf-schema#> という指定では、<http://www.w3.org/> 全体がバリデート・型変換の対象外になる。そのため、purl.org, [w3c.org](http://www.w3c.org/) 等を指定した場合、殆どの語彙がバリデート/型変換の対象外になることに留意する。
- ・一般的な外部の語彙の RDFS の管理が必要である (管理サーバ：
/var/www/html/lodmanage/commonschema/ 以下)。ただし、希に外部の語彙の RDFS のフォーマットに問題があることがあるため、一つづつファイルを導入し、都度データの変換を行って正常に変換出来ることを確認すること。

1.1.5. ソフトウェア配置

表 2-4-3-2-1 公開サーバ ソフトウェア配置

項番	ディレクトリ/ファイル	オーナー	chmod	用途	備考
1	/usr/lib/fuseki	loduser	755	fuseki インストールディレクトリ	
2	/usr/lib/fuseki/DB2	loduser	775	データベース(TDB)データディレクトリ	
3	/home/loduser/jrrktool/	loduser	777	公開サーバデータ更新インストールディレクトリ	
4	/var/log/fuseki/	root	777	fuseki ログディレクトリ	
5	/var/log/jrrk/	root	777	公開サーバデータ更新ログディレクトリ	
6	/var/log/httpd	root	777	Apache ログディレクトリ	
7	/var/www/html/	root	777	Apache 静的コンテンツ格納ディレクトリ	Apache のドキュメントルート
8	/var/www/html/rdf/	loduser	777	RDF 配信ディレクトリ	
9	/var/www/html/schema/	loduser	777	RDF スキーマ情報ファイル配信ディレクトリ	
10	/etc/httpd/conf.d/	root	755	Apache 個別設定ファイルディレクトリ	
11	/etc/httpd/conf.d/fuseki.conf	loduser	644	Apache 個別設定ファイル (公開サーバ設定)	

表 2-4-3-2-2 管理サーバ ソフトウェア配置

項番	ディレクトリ/ファイル	オーナー	chmod	用途	備考
1	/home/adminuser/jrrktool/	adminuser	777	データ登録ツール (CSV/SDF→RDF 変換、 バリデータ、オンライン処理スクリプト、バッチ処理)	

				スクリプト、管理サーバ更新データ一覧) インストールディレクトリ	
2	/home/adminuser/jrrktool/jar/	adminuser	755	データ登録ツール jar ファイルディレクトリ	このディレクトリ中にバックアップファイルを置かないこと。
3	/var/log/jrrk/	root	777	データ登録ツールログディレクトリ	
4	/var/log/httpd	root	777	Apache ログディレクトリ	
5	/var/www/html/	root	777	Apache 静的コンテンツ格納ディレクトリ	Apache のドキュメントルート
6	/var/www/html/rdf/	adminuser	777	公開サーバ向け RDF 配信ディレクトリ	
7	/var/www/html/schema/	adminuser	777	公開サーバ向け RDF スキーマ情報ファイル配信ディレクトリ	
8	/var/www/html/lodmanage/	adminuser	777	データ登録ツール -UI, Upload インストール先	Apache で、このディレクトリ以下に Basic 認証をかける。
9	/var/www/html/lodmanage/style/	adminuser	755	スタイルシート、エラーファイル表示用 xslt ディレクトリ	
10	/var/www/html/lodmanage/commonschema	adminuser	777	共通 RDF スキーマディレクトリ	
11	/var/www/html/lodmanage/failed	adminuser	777	データ登録ツール変換・バリデーション結果エラーファイルディレクトリ	
12	/var/www/html/lodmanage/ignorelist	adminuser	777	データ登録バリデーション対象ボキャブラリ URI リストファイルディレクトリ	このディレクトリ中にバックアップファイルを置かないこと。
13	/var/www/html/lodmanage/inc	adminuser	777	データ登録ツール UI 共通ヘッダ、プッダファイルディレクトリ	
14	/var/www/html/lodmanage/syncupload	adminuser	777	Upload によるオンライン処理のデータファイルのアップロード先ディレクトリ	
15	/var/www/html/lodmanage/tempout	adminuser	777	データ変換終了・バリデーション前のファイルのテンポラリ出力先	

16	/var/www/html/lodmanage/upload	adminuser	777	Upload によるバッチ処理 のデータファイルのアップ ロード先ディレクトリ	
17	/var/www/html/listrdf.php	adminuser	777	管理サーバ更新データ 一覧：RDF の一覧表示	
18	/var/www/html/listschema.php	adminuser	777	管理サーバ更新データ 一覧：RDF スキーマファイ ルの一覧表示	
19	/etc/httpd/conf.d/	root	755	Apache 個別設定ファイル ディレクトリ	
20	/etc/httpd/conf.d/manage.conf	adminuser	644	Apache 個別設定ファイル (管理サーバ設定)	
21	/etc/httpd/conf.d/php.conf	adminuser	644	Apache 個別設定ファイル (管理サーバ PHP 設定)	

以上