

## CSVファイルの取り扱いに関する機能の作成に対して取り組んだ内容

このシステムでは、大量の情報を一括で投入したり、保存済みのデータをCSVファイルとして出力することで表計算ソフトなどで集計できるようにする機能を設けています。

この資料の目的は、CSVファイルの取り扱いプログラムで工夫したことをプレゼンする事です。

### 目次

1

選定したライブラリ

2

受付可能なCSVファイル

なお、取り込んだCSVファイルのデータは、メモリリークを避けるため一旦一時ファイルにCSVファイルとして出力をします。その一時ファイルを、他のクラスで共有して使用します。

詳しくは、別途資料「大きなデータの取り扱いの説明」に記述してありますので、是非ご覧ください。

## 選定したライブラリ

選定したのは、**Apache Commons CSV**となります。

理由としては以下のようになります。

### > 幅広い形式のCSVファイルに対応するため

CSVファイルとは、大まかに言えばテキストデータをカンマ区切りしただけの文章であり、Javaの標準の機能を用いて、カンマ区切りでテキストデータを切り分ければ、簡単なCSVファイルは抽出可能です。

ですが、現実には単純なものではなく、様々な形式のデータが存在し上手くいかない場合があります。

#### ・データの中にカンマが入っている

(abcd,efgh,ijkl,m nop)

切り分けて  
いいカンマ

切り分けてはダメなカンマ

#### ・データの中に改行が入っている

(あいうえお,かきくけこ,さしす~~¥~~nせそ) など

バグが入る確率を減らすためには、ライブラリを用いた方が  
確実と判断しました。

### > 信頼性のあるライブラリのため

CSVファイルの為のライブラリは、検索してみたら数多く存在しました。ですが、一部のライブラリは、最終更新日が結構前だったり適切な管理がされていない物でした。

今回選んだ「Apache Commons」は、頻繁にアップデートがされており適切に管理されていると判断した為、採用しても大きなバグを作りこむ恐れは少ないと判断しました。

## 受付可能なCSVファイル

このシステムでは、以下の全ての条件に当てはまらないCSVファイルは  
**全て受け付けない**ようにしております。

これは、不正なファイルのデータを取り込んでしまい、システムの  
他の部分への影響を防ぐためです。

(例：正常なデータ)

管理番号	名前	ふりがな	担当楽器
abc-123	森本	もりもと	トランペット
efg-456	伊藤	いとう	ヴァイオリン
hij-789	佐藤	さとう	フルート

・必要な列が欠落していない事

管理番号	名前	担当楽器
abc-123	森本	トランペット
efg-456	伊藤	ヴァイオリン
hij-789	佐藤	フルート

・ヘッダー行の文字列に誤りが無い事

管理番号	名前	まちがいの	担当楽器
abc-123	森本	もりもと	トランペット
efg-456	伊藤	いとう	ヴァイオリン
hij-789	佐藤	さとう	フルート

・列が所定の順番通りに並んでいる事

管理番号	名前	担当楽器	ふりがな
abc-123	森本	トランペット	もりもと
efg-456	伊藤	ヴァイオリン	いとう
hij-789	佐藤	フルート	さとう

・余計な列が存在しない事

管理番号	名前	ふりがな	担当楽器	余計な列
abc-123	森本	もりもと	トランペット	いない
efg-456	伊藤	いとう	ヴァイオリン	いない
hij-789	佐藤	さとう	フルート	いない

・行数が制限数を超えていない事

管理番号	名前	ふりがな	担当楽器
abc-123	森本	もりもと	トランペット
efg-456	伊藤	いとう	ヴァイオリン
hij-789	佐藤	さとう	フルート
asda	ads	fhsd	jrtu
asd	sda	fasd	dsfgds
ads	ad	asda	gfsdgs
ad	ga	fgse	sdfas
sd	arw	dsfad	asfdas

・データの中に、極端に長い文字列が存在しない事

管理番号	名前	ふりがな	担当楽器
abc-123	森本	もりもと	あい j h f さ j h t w h k l d h s f k l d f h d k j は い k d f h か l j h d f k あ h d k f じ ゃ h f k l j d f h k あ d f h か h d f あ k j h f k あ い j h f さ j h t w h k l d h s f k l d f h d k j は い k d f h か l j h d f k あ h d k f じ ゃ h f k l j d f h k あ d f h か h d f あ k j h f k あ い j h f さ j h t w h k l d h s f k l d f h d k j は い k d f h か l j h d f k あ
efg-456	伊藤	いとう	ヴァイオリン
hij-789	佐藤	さとう	フルート