

先月門野さんの発表で、「仕事は原稿整理が50%、InDesignでの組版が50%」というお話がありました。組版するまでの下準備がとても大切だというのは私も同感です。

今回の発表では、私が日常業務で実践している原稿整理法をご紹介します。

主にエクセルで入稿された原稿の取り扱いについてお話しします。

InDesignへの展開を前提とした場合、そのまま使えるエクセルのデータはほとんどありません。どうしても手作業が発生します。

私は、これまで定期的に入稿があり、量もそこそこあるものは、加工をプログラム言語で処理してきました。最近では、『Python』という言語で、機会学習のエンジニアがデータの前処理をする際に多用するライブラリ『Pandas』を使っています。

このライブラリについては、最近学習を始めたところでして、初歩的なところしか理解できていませんが、実際の仕事にどれくらい役に立つのか、サンプルを作りましたのでご覧ください。

（「学会リスト.numbers」を開く）このようなエクセルを題材としてどんなことが出来るのか見てみたいと思います。InDesignへ展開する表組の仕様は、

1. 列は「発表日」と「発表年」は不要なので削除する。
2. 列の順番は、演題 => 演者（または出演者） => 対談 => 学会名 => 開催初日 => 開催終了日 => 会場 => 開催地に変更する。
3. 『開催初日』と『開催終了日』を『開催初日-開催終了日』と『enダッシュ』で繋ぎ1列にまとめる。
4. 『判定』の指示に従って国内、国外で表組みを分ける。
5. 『演者』は7文字揃えにする。
6. 『対談』に『本人』とあるところは演者の氏名を4文字揃えで置き換える。

以上の仕様にプラスして、エクセルのセル内の改行は、InDesignの表組で展開すると行列が崩れますので、改行をとり去り『▼』でマークアップをする、というものです。

sample 1-1

では、1と2をやってみます。
列の入れ替えです。

- まず、この一行でセル内改行や日本語の正規化の処理をしています。
- そして、書き出す際に、演題 => 演者（または出演者） => 対談 => 学会名 => 開催初日 => 開催終了日 => 会場 => 開催地を書くだけ。

```
$ python sample1.py
```

sample 1-2

次に、3と4をやってみます。

- 『開催初日』と『開催終了日』の書式「月/日/年」を「年/月/日」に変換する。
- 『開催初日』と『開催終了日』を『開催初日-開催終了日』と『enダッシュ』を入れてまとめる。
- 『判定』の指示に従って国内、国外で表組みを分ける。

```
$ python sample1.py
```

sample 1-3

そして、5と6です。

- 『演者』は7文字揃えにする。
- 『対談』に『本人』とあるところは演者の氏名に置き換える。

```
$ python sample1.py
```

以上のように、Pythonを使うことで、列の入れ替え、行の抽出、文字列の正規化や任意の文字列の置き換えと柔軟な原稿整理ができます。

この柔軟さを活かして、洛北高校のPTA会報をやっていた時に思いついた、データ結合を使ってInDesignに配置させる顔写真の整理法をご紹介します。

InDesignのデータ結合用CSV書き出しに、

- 原稿に掲載されている氏名と関連付いた画像のファイルを検索し、
- ヒットしたファイルを任意の名称に置き換え、
- 置き換えたファイル名をCSVにリストする。

というものです。

「_org」フォルダーを開いておく

「+_教職員紹介.numbers」を開いておく

写真はUnicode順となり、名簿の順番ではありません。

これらのファイル名を変更しデータ結合用のCSVを作ってみましょう。

```
コマンド python sample2.py
```

確認してみましょう。

氏名、ファイル名。。。きちんとリストされます。

写真のファイル名も置き換わりました。

折角なので、これと同じデータがこちらにありますので、実際にInDesignのデータ結合を試してみたいと思います。
ルビはJavaScriptで自動化できる仕様に合わせてPythonから書き出しています。

最後に、これは Python-docx というライブラリを使って作成しました。

ワードデータの太字、イタリック、下線、上付、下付、およびそれらの複合した書式を色分けするコードです。
複数ファイルにも対応しています。

では、やってみましょう。

```
コマンド python yonde.py
```

これでもう色分けの暗記をする必要はなくなりそうです。

日報を見ると、5月の20日、研友会の顔写真付名刺の仕事からコードを書いたの原稿整理を再開しました。
まだまだ初級者ですが、これからも業務の役にたつコードを考案していこうと思っております。

ご清聴ありがとうございました。

