

練習問題・HTML生成ファイナル



ある依頼

ボス「キミ、わたしが読書家として評判なのを知っているかね」

あなた「どうしたんです、なにかご用ですか」

ボス「まあこれを見たまえ。わたしの蔵書リストだ」

mybooks.txt

あなた「おお。...これは本当に実在の本ですか。どうも正気でないタイトルがいっぱいありますが。著者の名前もどうも人間離れしているような」

ボス「そこは気にするな」

あなた「で、これがどうしたんです」

ボス「最近、友人のR君に何冊か貸したのだよ。R君自身がさっき、借りている本の覚え書きだということでこんなメモを残していった」

lendingbooks.txt

あなた「で？」

ボス「R君の喜びようを見ていたら、私の蔵書をもっと世に知らしめたいと思うようになったのだ。ついては私の蔵書リストをホームページで公開したいと思う」

あなた「ボス、自分のサイトなんか持ってるんですか。コンピューター苦手でしょ？」

ボス「うちの会社のサーバーにちょっと間借りすればいいんじゃないかな」

あなた「...まあ、いいんじゃないでしょうか。でもアレですよ、ウェブに載せるんなら、HTMLを書けなきゃダメですよ」

ボス「にやり。つまりそういうことだよ、キミ」

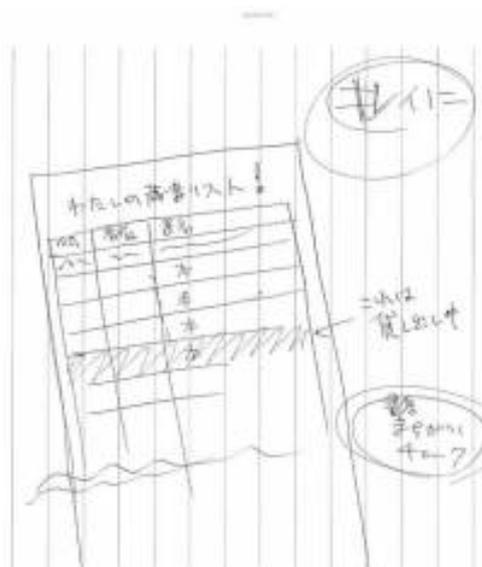
...

ということで、ボスの私用につきあって蔵書リストのHTMLを作るというのが今回の仕事です。

ボスいわく、「現在貸し出し中の本については、それが分かるようにマスに色なんかがついているといいなあ」だそうです。うわめんどくさ

どうやってやりましょうかね

ってことで、やってみましょう。その前に、ボスが書き残していった「仕様書」らしきものを解読してみましょう。



(クリックで拡大)

何が「キレイに」だまあ、意を汲んで、適当に整形してあげることにはしましょう。

で、「わたしの蔵書リスト」というのが表のタイトルのようですね。

表の中身は、どうもボス自身がつけている蔵書番号と、著者名と、本のタイトルがそれぞれ並ぶもののようです。ときどき行に色がついているのは、この本は現在貸出中と言いたいみたいです。色は、適当に決めておきましょう。

メモの端に「書名まちがいチェック」と書いてありますね。これは何だろう、と思いましたが、あなたはボスに聞きなおしてみました。

あなた「さっきの仕様で、『書名まちがいチェック』ってメモしてありましたが、あれは何がしたいんですか」

ボス「ああ。R君が借りている本の一覧には、書名も書いてあったら。たぶん、あっちが正しい」

あなた「？」

ボス「私自身の作った蔵書リストは、どうも時々タイトルを転記し間違っているようなのだ。せっかくなので、R君が書いた書名と照らし合わせて、私が間違っていたらあとで直しておきたい」

あなた「Rさんが間違っている可能性はないんですか」

ボス「まず、ない。とにかく、タイトルの不一致があったら、作るリストのどこかに適当な目印を打っておいてくれ」

そういうことだそうです。じゃあ、リストには「このタイトルはちょっと怪しいかも」ということを示す目印も、行ごとに必要なら打っておくことにしましょう。

あとは、ヒントとか

問題としては、ここまでの記述でほぼ全部説明してしまいました。あとはこれをどうスクリプトに落としていこうかな、ということについて、手当たり次第にヒントらしいものを述べていきましょう。

まず、今回、ひとつのスクリプトで二つのデータファイルを扱うことになりますね。ひとつはボスの蔵書リスト、もうひとつはR氏への貸し出し中リスト。ふたつ読み込むわけですが、まず大枠としては下のようなりかたでよいでしょう。

```
infile = open("first_data.txt")
for line in infile:
    何か意味のある処理...
infile.close()

infile = open("second_data.txt")
for line in infile:
    何か別の意味のある処理...
infile.close()
```

ファイルをopenするときの目印（変数）を別々のものにしてもいいですが、まあ、一回読み終わってcloseしたら、同じ名前でも使いまわしても別に紛らわしくないでしょう。このときは、データ書き出しのときと同じように、いったんファイルの目印（今後、時々ハンドラと呼ぶかも）を明示的にcloseしてから次のopenをするべきですね。closeを抜いてもそれなりに動作するんですが、ここらへんはマナー？みたいなもので。

で、どちらのファイルを先に読み込んでおくべきか、という計画になりますが... 貸し出し中リストを先にするほうが今回の仕事では自然じゃないかな、というのが、筆者の考えです。つまり、

- まずは貸し出し中リストを読み込みながらどっかに記憶しておいて、あとで蔵書リストを一行づつ読み込むときに「こいつは貸し出し中かな」という判断をその場でしたい

ですからね。

その際、貸し出し中リストは、辞書として読み込んでおくとういいます。つまり、貸し出し中リストを一行読み込んで（蔵書番号、著者名、タイトル）に切り分けたら、蔵書番号がキーになって、著者名とタイトルがそれに対応する値になるような辞書を構築していくということ。

こうすると、あとになって、「これって貸し出し中？」という問いにすぐ答えることができます
んで。

has_key という命令をまだ説明してなかったでしたっけ。これを覚えると、ある辞書に特定のキーが入っているかをすぐ確認することができますよ。下の対話シェル例のとおり。

```
>>> d = {'AAA': 'Book A', 'BBB': 'Book B'}
>>> d.has_key('AAA')
True
>>> d.has_key('CCC')
False
```

dという変数に辞書が入っていたとして、そいつにhas_key(何か)という指定をすると、そのキーがあるときはTrue, ないときはFalseが帰ってきます。だからif文をつかった条件分岐にもってこいですね。

(ところで、今まで無造作に「命令、命令」と言ってきましたが、d.has_key みたいに、「.」で始まって何かを指定するときは、今後は特に「メソッド」と呼ぼうかと思えます。今までにも、文字列について、.find とか、リストについて .append とかありましたね。あれも正確に語ろうとするならメソッドと呼ぶのが正しかったんです。ま、たかだか用語の問題ですけど)

貸し出し中リストの話はもうちょっと続きます。キーは「蔵書番号」ということでよいでしょう。その値として、著者名とタイトルを同時に詰め込みたいですね。こういうときは、あるキーに対して「リスト」を格納してしまうことにすれば解決です。

- 'B001' というキー → 値は、['著者A', 'タイトルA']というリスト
- 'B002' というキー → 値は、['著者B', 'タイトルB']というリスト
- ...

といった感じ。こいつも対話シェルから試しておきましょうか。

```
>>> d = {}
>>> d['B001'] = ['authorA', 'TitleA'] ←あるキーに対する値として、リストを突っ込んでおく
>>> d['B001']
['authorA', 'TitleA']
>>> d['B001'][0] ←キーを指定して、さらにそこからリストのように「何番目」の指定をする
authorA
>>> d['B001'][1]
titleA
>>> t = d['B001'] ←いったん別の変数にキーに対する値を取り分けておいてから、
>>> t[0] ←それに対して改めてリストのようにアクセスしてもいいよ
authorA
>>> t[1]
titleA
```

辞書的な処理とリスト的な処理が入り乱れていますが、大丈夫かな。どうかな。

続・ヒント

つぎに、HTMLを作っていくときにどんな感じの工夫が必要かな、という点についても若干。

いろいろとボスから注文がありましたね。下のようにならまとめみましょう。

- 現在貸し出し中でないもの。これについてはあまり問題はない。

- 現在貸し出し中。このときは、今は借りられないよ、ということを書き加えるために、セルの背景にちょっとした色をつけておきたい。
 - ただし追加仕様。貸し出し中で、しかもR氏のメモと本来の蔵書データで「タイトルの食い違い」があるときは、もうちょっと表示に工夫が必要。

タイトルに食い違い、と言いましたが、実はデータには著者名も含まれています。こいつもついでにチェックしてあげるといいかもしれませんね。

で、その食い違いがあったときはどういう表示をしてあげようか。

いくつかアイデアは挙がるでしょうが、安易ながら有効だと考えられるのは、R氏のデータのほうを優先して表示してあげることでしょうか。ただしそのときは文字の色を変えておくといいですね。ボス自身がこのリストを見てアツと気づかなくちゃいけないわけですから。まあ、この他に面白いやりかたがあるか、それは各自におまかせします。

もうちょっとヒントくれ、という話が出たらまた考えますが、まずはここまででやってみましょうか。

練習問題05

【練習問題：05】 ボスの蔵書リストをHTMLファイルで作成するスクリプトを作成せよ。蔵書データは `mybooks.txt` である。ただし、R氏への貸し出し中データ `lendingbooks.txt` も、本文中の指定に従って適宜加味せよ。ふたつのデータファイルはそれぞれシフトJISエンコード、タブ区切り形式である。

※データファイルは、右クリックから「リンク先を保存」「対象を保存」などを選んで保存してください。