

# 【補講】新着の自動通知

## 準備のお願い

空き時間に、Google アカウントにサインインした状態で以下の2つのファイルをコピーして、自分専用に別名で保存しておいて下さい。

その後は必ず、元のファイルも、コピーしたほうも、閉じておいて下さい。

**[bit.ly/JNPC1110kaiken](https://bit.ly/JNPC1110kaiken)**

と

**[bit.ly/JNPC1110issei](https://bit.ly/JNPC1110issei)**

です。上がサイト更新の自動通知、下がチームでの一斉作業の練習です。

本日の講義終了後は、コピーできなくなります。

お問い合わせ、後日のフォローなどは、ご遠慮なく

[yoshito.nishio+JNPC@gmail.com](mailto:yoshito.nishio+JNPC@gmail.com)

までお願いします。

## ■サイトの新着情報を自動通知する

たとえば、日本記者クラブの会見予定のページ。定期巡回して、会見予定が追加されたり、時間が変更になっていたら、通知するメールを自分あてに送る。Google Apps Script を使う。

いまサインインしている Google アカウントのメールアドレスが発信元&受信先になる。to や cc、bcc の受信先を増やすこともできる。URLを変えて、別のサイトをウォッチする下敷きにも使えるが、その場合はコードの手直しが必要。

以下のスプレッドシートを本日開講中だけコピー可能にしてある。

**[bit.ly/JNPC1110kaiken](https://bit.ly/JNPC1110kaiken)**

① 「ファイル」→「コピーを作成」。そのコピーを好きな名前に変更する

- ② 「ツール」→「スクリプトエディタ」を選んで、コードを開く

コード.gsという名前のプログラムの画面になる。中は、いくつかの「ファンクション」に分かれている。使うときにはメインの「autoMail」を起動すること。ほかは、このメインプログラムの言い付けを実行する手下たち。

- ③ 左上からやや真ん中に寄ったところにある「関数を選択」で「autoMail」を選ぶ

- ④ 左の方にある「再生ボタン (▶)」みたいな右向き三角マークを押す

承認手続き※1がポップアップする。ここでやっておかないで後回しにすると、Chrome のブロッカーが悪さをして、トリガー設定時にエラーすることが多い。

- ⑤ その左のほうにある「時計」マークを押して、トリガーを設定

と進む。トリガーは、田中さんのマニュアルにもある通りで

- ⑥ 右下の「トリガーの追加」を押す

- ⑦ 開いた窓の3番目、「イベントのソースを選択」を「時間主導型」に

- ⑧ デフォルトでは1時間に1回、巡回する。そのままでもOKだが、頻繁にチェックさせたいければ、「分」で「10分」などに適宜変える。無料版の場合、メールは1日に計100の宛先にしか送れない、という制限があるので、よく考えて

#### ※1

「承認が必要です」とか「このアプリは確認が済んでいません」と出たら、承認、または先ほどコピーを別名保存したときの名前の「××××のプロジェクト（安全ではないページに移動）」を押す。田中さんのマニュアルを参照のこと。「安全なページに戻る」のボタンしかない場合は、「詳細」を押す。「次へ」「承諾」と進むと、自分で自分にメールを送る、という仕組みが動かせるようになる

#### ※2

Google アカウントを複数持っていて、両方でログインしているような場合（Gmailのアドレス2つ以上を使い分けて、タブを同時に複数開いてメールを読んでいる、とか）ような場合、スクリ

プトが起動できずエラーすることがある。いったんログアウトしてから、スプレッドシートのオーナー（編集可能な持ち主）になっているアカウントでログインし直すとよい

プログラムでやっていることは、//や/\*以下のコメント（プログラム中にあるが、コンピューターは実行しない部分。人間のためのメモ書き）に入れているので、興味のある人だけでも、億劫がらずに眺めてみて。

（あ）巡回先のURLや、メール送信先、文面などの初期設定をする

（い）指定したページのHTMLを取得して保存する

（う）前回取得したときのデータと違いがあるか比較する

（え）違いがあれば、次回との比較に備えて、データをシートに保存する。  
→違いがなければ、何もしない

（お）さらに、HTMLコードの差分が、本当に会見の新着なのか、そうではないのかを調べる。  
→新着でなければ、これ以上は何もしない

（か）会見の新着を知らせるメールを送る

という流れ。日本記者クラブの会見予定の場合は必須だが、サイトによっては（お）の手順は不要な場合もある。

ところで、「HTMLに差分があるのに、新着ではない」というのはどんなとき？

## ■チーム作業でデータを集める

たとえば、議員の公式ホームページやSNSの一覧を作りたい、という想定。取材チーム員で手分けして、シートにデータを足していく、という作業を試みる。

構成のしっかりしている、自民党のホームページを見てみるとよい。さて、どのページを手がかりにするのがよいだろうか。

**[bit.ly/JNPC1110issei](https://bit.ly/JNPC1110issei)**

をコピー可能にしてある。会場にいる全員で、このファイルにアクセスする。

① ファイルのコピーを保存。「ファイル」→「コピーを作成」。

これは、ひな形として持ち帰っていただくため。

② コピーは閉じて、元ファイルのほうを開く。会場にいる全員で、同時に触っている状態

③ 自民党のページを開く。方針を考える

- ・どのページからデータを取るのがよいか？
- ・一度で必要な項目を取得できるか？ 2ステップ以上必要か？
- ・ツイッターをやっているかいらないか、など、ばらつきがあるようだ。もれなく取得するために、手分けして必要な項目を列挙していく。公式ホームページ、ツイッター、フェイスブック……
- ・スプレッドシートに仕事させる「式の入力」でラクをする手はないか？
- ・XPathの指定はどうしたらよいか？

XPathについては、副島さん、田中さんの講義を参考に。

#### 【余談】XPath Helper

Chrome の追加ツール「XPath Helper」があると便利。これで試してみて、欲しいものが取れていれば、だいたい大丈夫。長いXPathのままで可、無理に手を入れなくてよい。うまく取れないときには、HTMLのソースを見ながら試行錯誤する。

「id」は一意なので、これが出てきたら、そこから先の枝だけ探させればよい。つまり、根元や幹から丁寧に書かずに

```
//*[@id='xxxxxxxx']
```

と、それに続く部分だけに削って短くしても、動作する。

#### 【余談】絶対参照

シートに式を入力するとき、「F4」キーを押すと、セル名の「\$」マークがついたり消えたりする。この\$がついたカラム（縦の列。英字で書く）や行（こちらは横。数字で書く）、あるいはその両方は、式をコピーして増殖させたときに、参照先がずれずに固定される。なので、固定したい部分は\$つき、気を利かせてスライドしてもらいたいなら\$なし、と使い分ける。

前回も紹介したとおり、Excelでも使える小技。とても役に立つので、これだけはぜひ、マスターして！

④ 「か行」……「や行」「わ」を分担して作業

こちらで、ひな形のシートをコピー（CTRL+ドラッグが早い。その後、シートの名前を直しておく）するので、テーブルごとに、担当分の作業を。

⑤ 赤の太枠のセルに、親となる式を入れる

先頭にアポストロフィーをつけて、式ではなく文字列として仮置きしてあるところもある。アポストロフィーを削って使ってもらってもよい。

その式を、下や右にコピーして増殖させる前提で、ラクする方法を考えること。

ヒント：\$マークで、固定とスライドを使い分けて

【余談】

複数の項目が検索条件にマッチしたとき、Google スプレッドシートでは、続くセルにどんどん書き出してくれる。これで助かるときと、かえって困る場合とある。

もし、1つだけ書き出すことにして、2つ目以降は捨てるなら

`array_constrain()`

を使う。

数式を1つのセルにだけ書いて、それを次々と下のセルに適用したいときは、式を入れ終わって確定するときに、「Enter」単独ではなく、「Shift + Ctrl (Macは⌘) + Enter」を押す。

または

`arrayformula()`

で式を囲んで使う。配列式、というものになる。

⑥ 親の式を、下や右に増殖させる。一度に1カラムずつぐらいつつ

⑦ データが取得できたカラムは、コピーして「値の貼り付け」(Shift + Ctrl + V) に直す

でないと、負荷が重すぎて、データの一部だけしか取得しきれない状態になってしまう

⑧ ④～⑦を繰り返して、全員分が埋まったら、完成

⑨ 必要であれば、各自で完成品をコピーして、別名で保存を。①と同じ手順

お疲れさまでした！