

How-to guide to using web automation tools to collate impact evidence from social media.

 blogs.lse.ac.uk/impactofsocialsciences/2014/08/08/web-automation-tools-social-media-ifttt-zapier/

8/8/2014

*With such diffuse streams of web activity and academic engagement, there is a great need for simple ways to capture and record this valuable data, whether for personal use or reporting purposes. **Alistair Brown** provides an overview on useful automation tools that streamline the process. Over time, depending on the scope of your search and social media activity, setting up these feedback loops can be a relatively straightforward way to collating all online discussions as a method of proving impact and relevance.*



Previously on this blog I have described [how to use Twitter](#) to track the real-time impact of a media dissemination event. As I discussed there, one problem with Twitter is that it does not maintain an easily searchable archive of tweets, meaning that any engagement activity may be lost if not captured more or less as it happens.

Whilst it is possible to collect such information for a one-off event, any researcher or department that develops impact or dissemination activities will be generating feedback from the public and other external bodies all of the time, via numerous channels such as Facebook, Twitter, blog comments, social bookmarks, newspaper columns and comments, and many more. How best, then, to capture this ongoing torrent of data, which will need to be collected over the next five years ahead of REF 2020? Collating feedback can be a very time consuming process, and whilst there are professional social media and marketing companies who will happily take on the work, there are cheaper alternatives – especially for smaller departments or individual researchers.



Image credit: [IFTTT – Tweet after blogging by Wesley Fryer](#) (Flickr, CC BY-SA)

Thanks to web automation tools such as [If This Then That](#) (IFTTT) and [Zapier](#), it is possible freely and automatically to capture potential dissemination and impact evidence, and to collate it into a single document that records the various places where engagement may have occurred. Rather than you needing to monitor your outreach sources

regularly, such a document can passively grow in the background until needed, and is not dependent upon the availability of the original source several years hence. Imagine the difficulty in proving engagement for your brilliant 2014 outreach activity if, as might conceivably happen, Twitter or Facebook decide in 2019 to prevent APIs from searching historically, or if the EU's "right to be forgotten" rule has erased important search results from the web.

Automation tools work by collecting data from a source at trigger points you specify (such as when you publish a new Facebook post, someone posts a tweet with particular keywords, or an RSS feed has a new item) coupled with a channel through which those results are represented or redistributed.

There are various potential uses for this – for instance, I have connected [READ's](#) Google event calendar to our [Twitter](#) and [Facebook](#) accounts so that at set intervals (such as one week before an event, and two hours before an event) it will tweet and post reminders to our followers. For impact purposes, though, Zapier is most powerful as a way of connecting your social media channels with a Google document into which all of the chatter from various sources can be funnelled. Setting this up is quite tricky, but has the enormous benefit of bringing together all the useful statistics and evidence of engagement into one place, in a reliable archive.

Image credit: [Google Docs Icons 512px + vector](#) by [carlosjj \(DeviantArt\)](#)

To do this, you will need to have a Google Docs account, and to setup an account on Zapier. Create a [new Google spreadsheet](#), and separate worksheets that correspond to each source that you want to capture. For example, in the same spreadsheet you might have a worksheet to capture discussions from Twitter, another worksheet for comments from Facebook, and a third for comments to your WordPress blog. In each worksheet, create a row of column headings with the information you need. When capturing data from Twitter you might want to know the username of the person commenting about you, his or her location, the date of the tweet, and the tweet content. When capturing WordPress comments you will probably want to know the name and location of the commentator, the comment, plus the URL of the post they have commented on and, if the comment is a pingback, the URL of the page they have linked from. If you want to trace the impact of articles published the mainstream media, you could use Zapier to capture the output of comments by drawing on the comment RSS feeds, filling in fields such as commentator name, date, and comment content.

- You then need to fill the next row of each worksheet with some sample data corresponding to each heading, so that Zapier knows how the spreadsheet is structured.
- Now logon to Zapier, and create a new [Zap](#) – a Zap being a name for a connection between the source account you are tracking and your output, in this case a document.
- Complete all the required information, which will vary depending on which data source you are tracking and how you want to use it.

For Twitter, for instance, you might want to use filtered search terms. For example, I set a search to document every Tweet containing the words Durham NOT North Carolina AND English OR Literature to try to capture activity relating to our department; I also have a Zap to capture every time someone posts to our Facebook timeline, and another Zap for comments on our WordPress [blog](#). You could set up separate spreadsheets for particular members of staff, one-off impact activities, or publications on which you want to trace feedback.

Over time, depending on the scope of your search and social media activity, you should have a comprehensive set of spreadsheets collating all online discussions of relevance. Depending on the complexity of your filters and connections, you are likely to collect some noisy results as well, so you will still need to sift through your data. Nevertheless, on balance, this approach to capturing impact activity has two key advantages: firstly, it automates an otherwise repetitive process; secondly, it ensures that you capture activity as it happens, and so are not dependent

on unreliable web archives when you need to gather information in five years' time.

Note: This article gives the views of the authors, and not the position of the Impact of Social Science blog, nor of the London School of Economics. Please review our [Comments Policy](#) if you have any concerns on posting a comment below.

About the Author

Alistair Brown is a Postdoctoral Teaching Assistant in the Department of English Studies at Durham University, where he edits the blog [Research in English At Durham \(READ\)](#). He is also an Associate Lecturer at the Open University, teaching courses on the Arts and Literature. He tweets as [@alibrown18](#) and blogs at www.thepequod.org.uk.

- Copyright © The Author (or The Authors) - Unless otherwise stated, this work is licensed under a Creative Commons Attribution Unported 3.0 License.