



Exploring the Twitterverse

DR ELLIOT MEADOR, SRUC

DR HENRY CREISSEN, SRUC

MS ABIGAIL WILLIAMS, SRUC

This research was funded by the Plant Health Centre and the SEFARI Gateway Responsive Opportunity Fund 2020. For more information visit <https://sefari.scot/sefari-gateway>

SRUC, West Mains Road, Edinburgh, EH9 3JG

Email. Elliot.Meador@sruc.ac.uk ; Henry.Creissen@sruc.ac.uk ; Abigail.Williams@sruc.ac.uk

Key Findings

This report reviews findings from two research projects that both investigated how online social networks are being leveraged to proliferate information related to sustainable agriculture in Scotland. Data was collected from Twitter for one year analysed using social network analysis and natural language processing. Results indicate that the following approaches may be useful in increasing beneficial engagement between farmers, scientists and agriculture stakeholders:

1. Scientist and agriculture stakeholders who want to share their research findings with the farming community should engage directly with farmers in a way that builds trust with them over time.
2. Scientists who want to promote their own research findings can identify advocates who will help them reach new audiences, called tweet launching.
3. Using images and videos of topical subjects will likely help improve the reach and online impact; farmers who share images of their farms and scientists who share publication-quality graphs of their research tend to have more reach for individual tweets.
4. Tweet often. Those who tweet more often tend to have higher number of followers and impact overall.

Introduction

Social media has the potential to allow for a diverse set of discussions between stakeholder groups working in the agriculture and rural sectors. The microblogging platform Twitter is often used by government and academic institutions to disseminate information related to new policies as well as scientific research with the aim of increasing impact and the uptake of new practice. The platform is also used at an individual researcher level, both as a personal and professional tool, to keep in touch with friends and family and to engage in community specific dialogue.

What is less known is the degree to which farmers and those engaged in the land-based economy use the platform to discuss topics related to agriculture and land use policy, climate change and ecological services. For the first time, we have identified and mapped the critical networks of farmers, researchers and policymakers in the Scottish rural sector.

Based on analysis of these networks, we have developed a **best-practice guide** on how to better engage with Scottish stakeholders via online social networking platforms, such as Twitter. Establishing criteria that can be used to increase engagement and impact online will benefit all professional users of social media and has the potential to increase the reach and scientific impact of the Scottish Government funded RESAS portfolio.

Methods

This best-practice guide has been developed based on research carried out by SRUC researchers and has been funded by the Plant Health Centre and SEFARI Gateway Responsive Opportunity Fund.

The research was carried out in three stages, with each stage exploring a different research agenda. The three stages were:

1. Data collection and social network mapping
2. Data analysis
3. Stakeholder workshop

Data Collection and social network mapping

This research relied on two primary methods of data collection and analysis. First, we developed a new method of collecting data from Twitter called *network jumping*. Network jumping allows us to collect data that more closely represents conversations on Twitter.

We also used an approach based on natural language processing that uses key-word phrases to classify Twitter users based on their biographical data. Users were classified into several stakeholder categories depending on how they identified themselves in their biographies (see Figure 1)

Figure 1: Twitter Profile example identifying key-word phrases used in natural language processing



Network jumping, shown in Figure 2 below, illustrates how the algorithm works. First a series of starting stakeholders is identified. The starting-stakeholder's tweets are downloaded daily. Each day, a computer programme searches through their tweets and identifies those users whom the starting stakeholders have conversations with – measured using Twitter mentions (@'s). Finally, these newly identified user's tweets are downloaded. The resulting database shows discussions between stakeholders' overtime and can grow quite large depending on the day.

Figure 2: Network jumping schematic, displaying how network jumping works to gather data.

Network jumping approach to gathering data from Twitter API

Network edges are drawn when one use mentions another in a tweet

For example: @Scottish-Policy: hey @Agronomist the weather is great!
draws and edge between @Scottish-Policy and @Agronomist

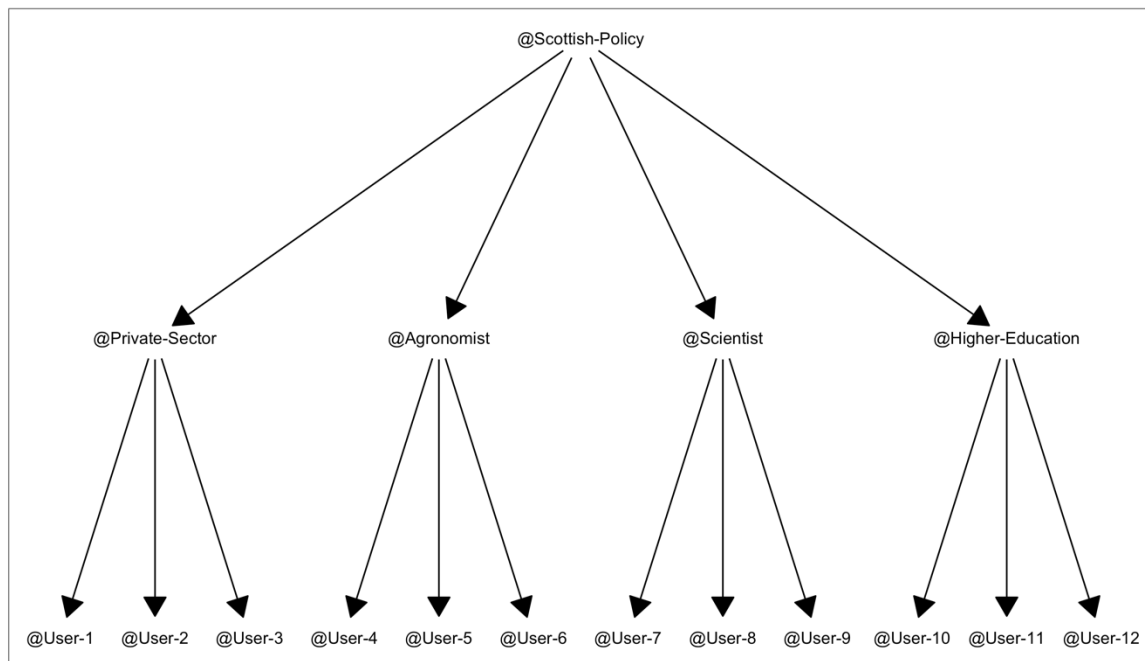


Diagram of the Twitter network jumping.

NOTE: Labels are randomly chosen and for illustrative purposes only.

Due to the network jumping approach to collecting data (with online conversations being the driving force behind who is included in the database) we can develop a social network by creating an edgelist of participants. An edgelist is a data formatting approach commonly used to create networks. We used the igraph (Csardi G, Nepusz, 2006) package in the R Programming Language to create the network objects and the ggraph package to do the majority of network plots. We used the Gephi software programme to plot the large network graph (Figure 4).

In total, we collected about 1 million tweets related to agriculture and rural society within Scotland over the period of 1 year (2019-2020).

Data analysis

For stage 2 of our research, using the same network dataset, we carried out some exploratory analysis and pulled out basic descriptive statistics. Some examples of data we identified was:

- Accounts that were tweeting most
- Accounts that had the highest number followers
- Accounts/tweets with highest numbers of retweets
- The most favourited accounts/tweets
- On what day, month and time do most accounts tweet
- Most popular #hashtags

In addition, several new variables were created within the dataset to reflect our need to quantify 'impact' and a new concept was developed called 'launching'.

Tweet impact

The tweet impact variable created within our dataset was based on whether a tweet had reached significantly beyond their own following. An impact score was calculated based on the number of times a tweet was retweeted in relation to their number of followers. Early preliminary analysis found that there was no statistically significant relationship in our dataset between the number of followers an account has and the number of times their tweet has been retweeted. This finding was important as it highlighted to us early within our research that there is potential for smaller accounts (such as individual researchers) to achieve similar levels of tweet engagement and impact as larger twitter accounts. By calculating tweet impact in a way that takes account of the number of followers of the original account, this can help to assess the 'relative' reach and impact a tweet has, rather than looking solely at number of retweets.

Tweet launching

The tweet 'launching' concept was developed to test how often twitter accounts with larger numbers of followers, were retweeting tweets from accounts with smaller numbers of followers. Our focus on identifying launchers generally looked at research institutions or organisations with thousands of followers, however the concept of launching can also apply to individuals with large twitter followings. Understanding and promoting tweet launching is important in helping researchers to increase their research reach via twitter and will help increase the scientific impact of the RESAS programme.

Two variables linked to launching were created for each retweet within the dataset. One variable contained a 'launch' score and the other a binary yes/no based on whether the retweet was considered launched or not.

Launch score was calculated as:

$$\text{Number of followers (retweet account)} / \text{Number of followers (original tweet account)}.$$

If a score above 1 was calculated, the retweet was considered launched. Twitter accounts overall launch scores were also calculated. This score was based on a combination of how often the account launched tweets and also how high the accounts individual retweet launch scores were.

Sentiment analysis

A sentiment analysis was carried out using the R package tidytext (Silge J, Robinson D, 2016). Sentiment analysis is a technique that allows researchers to assign emotional tags to text. It works by pairing the individual words in a block of text with emotional codes that have been previously developed by academics and experts in the field. Sentiment analysis does provide a good overview of the overall sentiment of a body of text, but it is not an absolute science. It often works best by pairing it with other Natural Language Processing techniques such as word counts, as we have done.

Stakeholder Workshop

After carrying out analysis of our dataset we held a stakeholder workshop with 20 participants to share and discuss our findings. These stakeholders included researchers from across SEFARI, agronomists and farmers. Using breakout rooms, stakeholders shared their experience in using twitter and other social media platforms to promote their work and helped us identify best practices which may help increase reach and impact of tweets.

Figure 3: Stakeholder workshop breakout session questions

Breakout questions

- **Do you feel limited in the types of things you can say or promote on Twitter?**
- **Do you feel *comfortable* promoting yourself or your organization?**
 - Do you feel the same with promoting others?
 - What might change the way you feel about promoting yourself or others?
- **Is it easier to engage with someone if you know them personally?**
- **Do the findings from the large stakeholder graph (see figure 4) reflect your understanding of Twitter as an engagement tool?**

That is, do you see

 - Highly dense clusters of stakeholders discussing with one another.
 - Some crossover between research and farmers
 - A few discussion “brokers” that help link producers and researchers.
- **Have you ever promoted or *launched* another user, or have you ever been promoted or *launched* by another user?**

If so,

 - Did this help promote your work?
 - Is this something that would be useful going forward (i.e., is it a reliable policy to actively promote other users?)

If not,

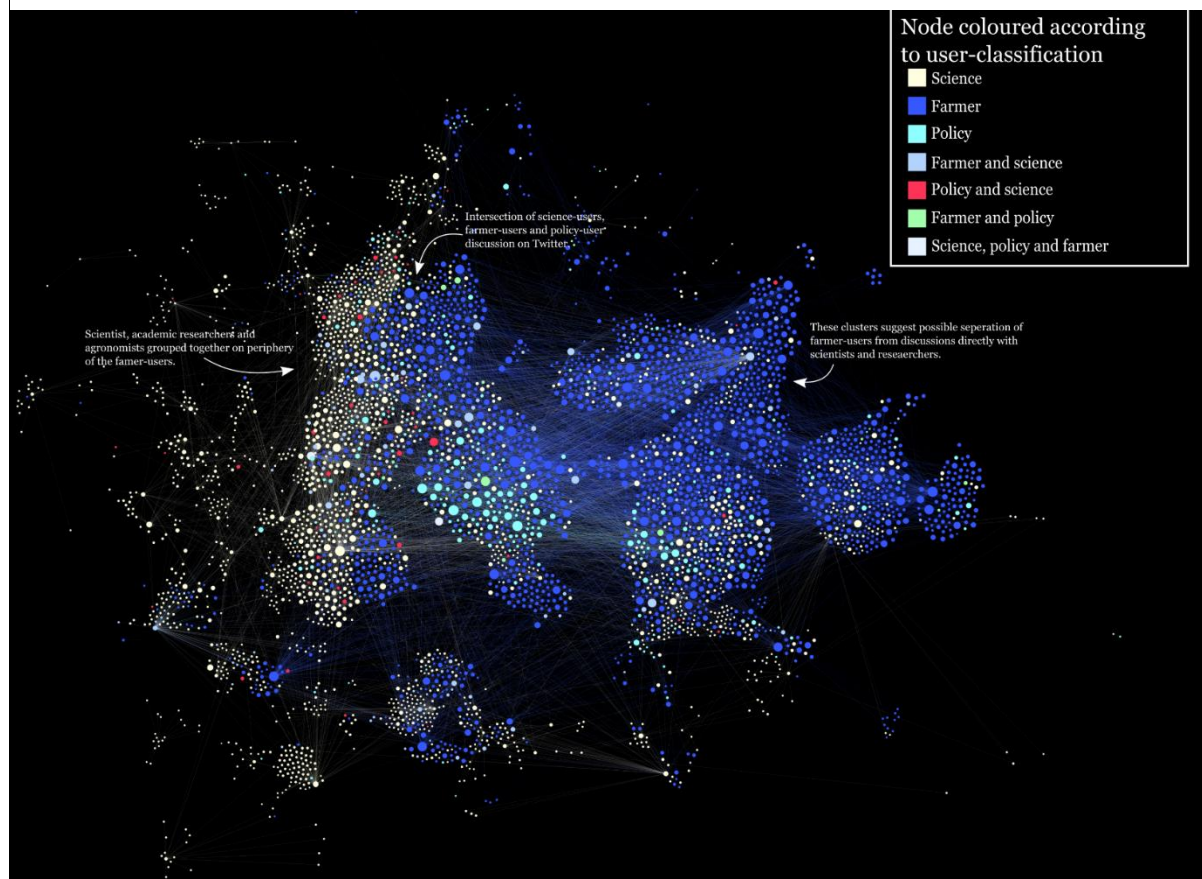
 - What are the primary barriers to being promoted by another user or institution?

Key findings

"Farmers tweet and they talk about Farming" - Using twitter data to confirm the 'obvious' and highlight the 'unknown'.

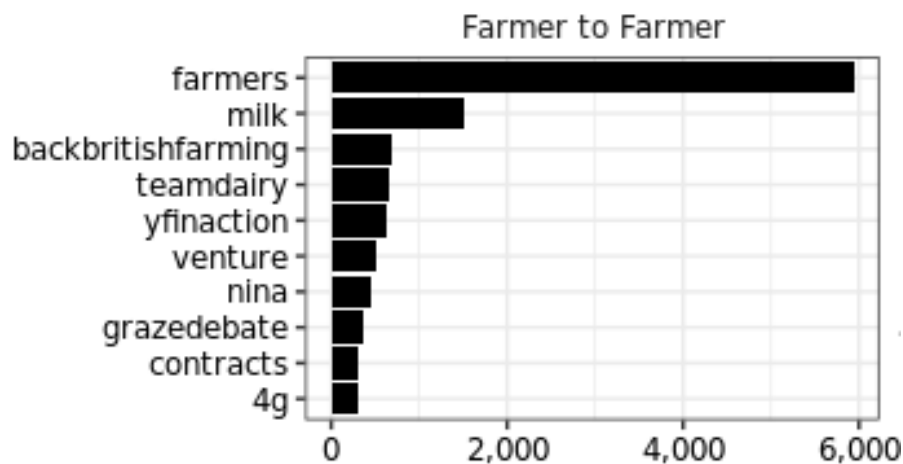
Figure 4 shows the overall network graph with each node coloured according to the category of each Twitter user. It shows primarily two stakeholder groups, Scientists (yellow) and Farmers (blue) engaging in discussions. We can see that each stakeholder group tends to have densely connected conversations with members of its own group. This is evidenced by the fact that there is not a lot of overlap between the yellow and blue points. However, there is a line on the left-hand side where many scientists are connected to farmers, indicating that some discussions are happening at these places in the network. Often, these discussions are being facilitated or promoted in some by other stakeholders, primarily policy/science stakeholders. This suggests that farmers and scientist's relationships online are benefited by network brokers, or those individuals that may help nurture online discussions between groups.

Figure 4: Twitter network of farmers, research scientists and policymakers



Our findings suggest that farmers are quite active on social media and (unsurprisingly) tweet about farming – 43.1% of users within our dataset are farmers and the most common word in farmer-to-farmer tweet interactions was ‘farmers’ (see Figure 5). Farmer networks on twitter appear to be quite strong and dense, however farmers tend to group with other farmers and research scientists tend to be on the periphery of these networks (see Figure 4). Research scientists working in agriculture and rural sectors who wish to increase their impact online should engage directly with farmers on twitter. In addition, the network map in figure 4 highlights opportunities to engage with well-connected individuals that can act as ‘brokers’ between stakeholder networks. Often, popular words in tweets are trending hashtags, or phrases meant to draw attention to a tweet (characterised by including the # symbol). This is illustrated in Figure 5, with the words “yfinaction” and “nina”, which are both hashtags related to agriculture.

Figure 5: Top words used in tweet interactions between farmers within the dataset



Increasing tweet reach and impact – assessing the potential for tweet launching.

Between 2019-2020, 49.87% of retweets within our dataset related to agriculture and rural society were classed as ‘launched’. However, Twitter account launch scores varied significantly between research institutions. Only two Scottish research institutions featured in the top ten twitter accounts for highest launch scores.¹ No Scottish research institution featured in the top five. Although there appears to be a culture of tweet launching among some twitter accounts, the vast majority of tweets (based on our tweet impact metric) are still ‘non-impactful’ (98%). In other words, the majority of tweets do not have a reach beyond their own twitter following.

¹ This also includes research institutions out with the SEFARI portfolio

Figure 4 visualizes the discussions occurring between different types of stakeholders on twitter. Figure 6 shows a similar network, although now focusing instead on tweet launching between stakeholders. A similar pattern can be seen to that of conversations between stakeholders (see Figure 4), where launching appears to occur most commonly between the same stakeholder categories. This is particularly clear in the case of agriculture – this can be seen in the yellow cluster in the bottom left of Figure 6. We see that science stakeholders are often disconnected from launching other stakeholders.

Figure 6: Network graph of twitter launchers between stakeholder category

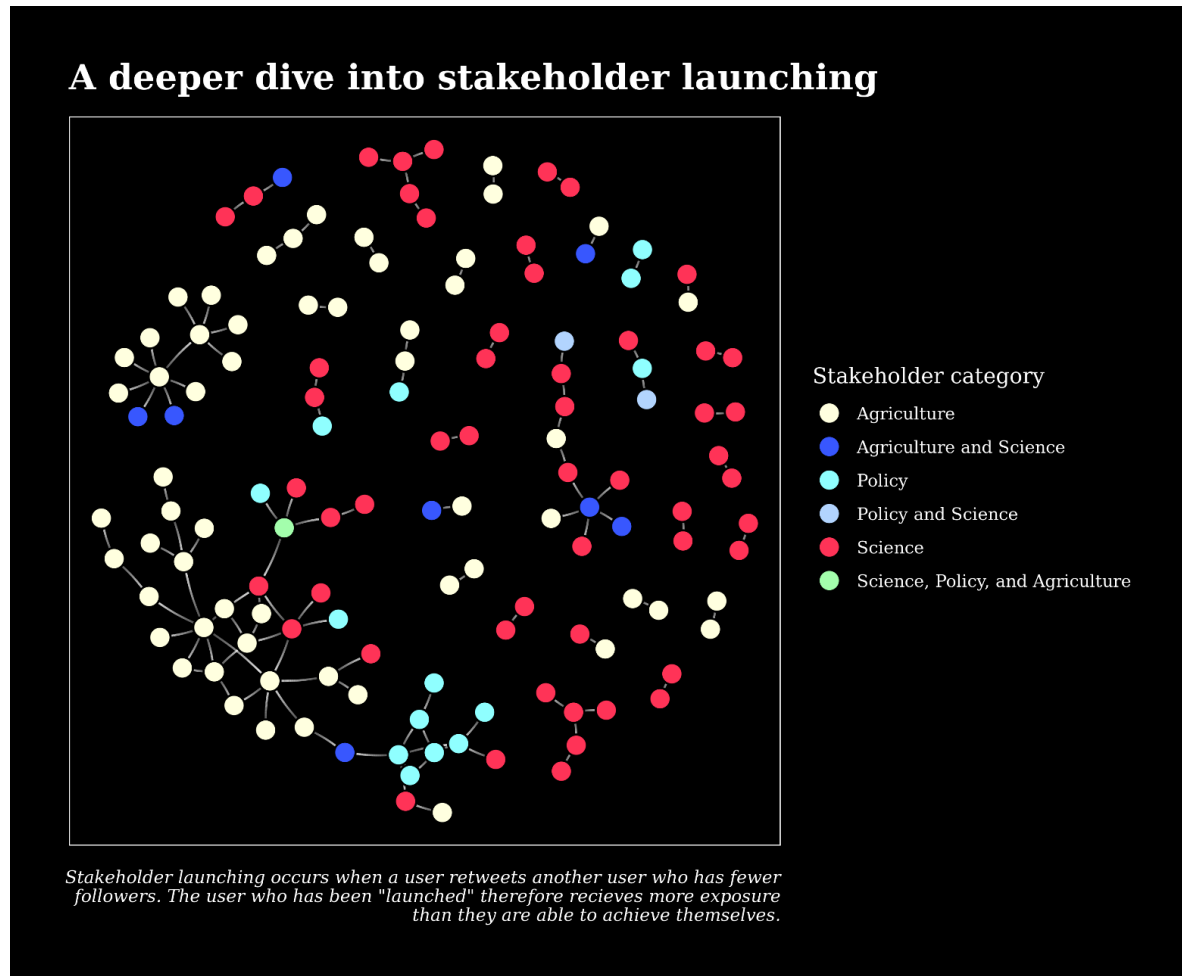


Figure 7 shows these launching patterns in more detail. What is clear from this table is that stakeholders that are identified as carrying more than one ‘hat’ i.e. ‘Farmer AND science’ or ‘Science, Policy AND Farmer’, are more likely to launch tweets from stakeholders not like themselves than stakeholders appearing from only one category. There is particularly low launching by scientists of tweets from farmers and policy.

Figure 7: Table describing % of launched tweets by stakeholders in different category from them

Original stakeholder	Launched	Percent (not like original)
Farmer and Science	Science	77%
Science, Policy, and Farmer	Science	54.50%
Science, Policy, and Farmer	Policy	45.50%
Policy	Farmer and Science	16%
Farmer	Science	12%
Science	Farmer	5%
Science	Policy	5%

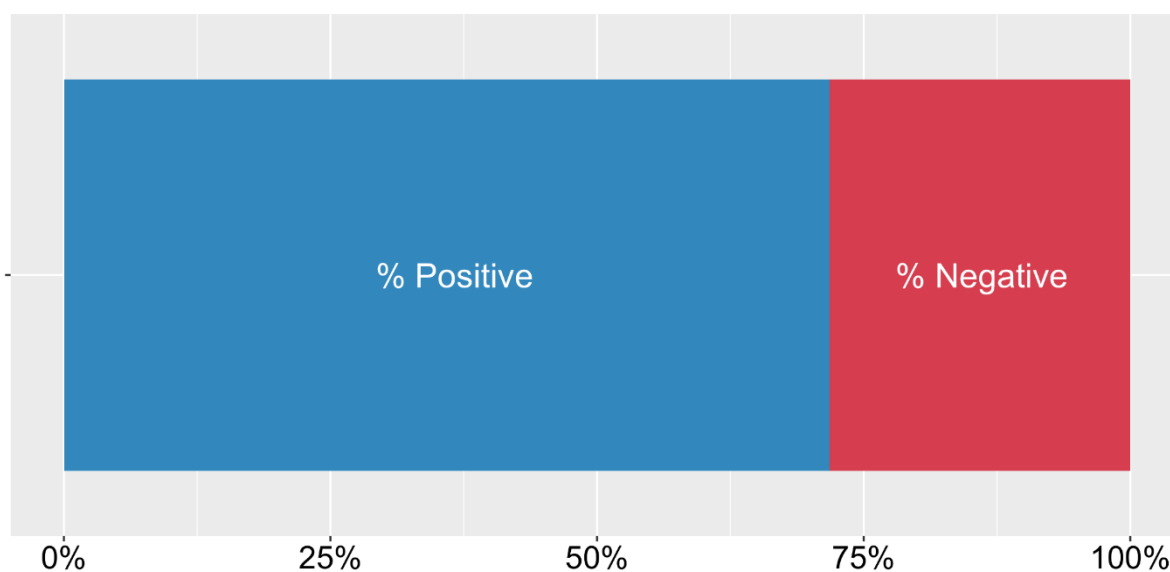
Navigating the Twitterverse – Maximising engagement through social media.

The project investigated how to increase research impact on Twitter by identifying the best ways to engage with stakeholders. Rules of engagement were developed during an online stakeholder workshop.

An exploratory sentiment analysis did not show any relationship between tweet sentiment and tweet 'impact'. However, stakeholders from the workshop did note that controversial topics were likely to be avoided during twitter exchanges – particularly from institution accounts. Workshop participants felt that sharing high quality evidence in the form of peer-reviewed scientific articles etc. is advisable when engaging in a controversial debate. This is also more likely to result in re-tweeting and less likely to initiate backlash. Our analysis showed that more than two thirds of tweets in our dataset had a positive sentiment (See figure 8) which may reflect comments from stakeholders during the workshop, about their reluctance to engage in negative interactions online. Suggestions were also made that having separate social media accounts for your personal and professional life is advisable to avoid any conflicts.

Figure 8: Percent of tweets in the dataset with a positive or negative sentiment score

Total Sentiment of Tweets



Suggestions to increase tweet engagement such as 'likes' and 'retweeting' included:

- Using media, such as videos and pictures, to draw attention to your Tweet
- Using topical hashtags
- Tagging institutions & individuals that may be interested in your tweet
- Using links to further information/research papers etc in your tweet

The potential for increasing the reach of your tweets by 'launching' was considered within the workshop. This approach seemed generally favourable by participants, although it was noted that information and support in identifying the most relevant tweet launchers to engage with may be necessary. Large institution accounts appeared to be a good option for tweet launching, with researchers already well connected to their own institution's communications and social media contacts. However, our data suggests that tweet launching is currently not a common practice within Scottish research institutions and some work may be required to engage with Communications and Social Media colleagues to highlight mutual benefits of the practice.

In terms of increasing launching between stakeholder groups, participants discussed the potential for increasing the number of farmers launching tweets from science stakeholders. Suggestions included engaging with farmers tweets online by liking and commenting on photos, sharing research that may be relevant to their interests and building mutual trust over time. This should be a two-way interaction, not one-way from science to farmer. Identifying stakeholders with multiple caps (e.g. Farmers AND Scientists) and asking them to launch tweets will also help to reach an audience beyond a single stakeholder group. These brokers play an important role in disseminating information between policy, science and farmers and engagement with them online could provide a quick and effective way for researchers in disseminating research beyond the scientific community.

Conclusion

Social media platforms such as Twitter offer users the ability to engage with a wide range of stakeholders and users. This ability does not come without costs, and engagement in online spaces related to scientific research can be daunting and sometimes overwhelming. Often this overwhelming feeling can induce some researchers to disengage in online spaces related to their work. Our research shows that there are a few practical approaches to engaging online that hopefully can make engaging feel less daunting and more enjoyable overall, which will hopefully lead to more impact in your research.

Framing your research findings in a positive way will be more beneficial than doing so in a negative light. Also, sharing peer-reviewed findings of one's work will help it get noticed and "liked" by larger institutions. Tweet launching is perhaps the best way to promote new research findings. To do this, it's best to engage with a few large institutional accounts prior to sharing results, this way these institutions will be waiting to help promote it (they may even offer feedback that allows them to share it). In the end, it's important to remember that (nearly) all users are people, and so engaging with them as they are people is the best way to build rapport with them.

References

Csardi G, Nepusz T: The igraph software package for complex network research, InterJournal, Complex Systems 1695. 2006. <https://igraph.org>

Silge J, Robinson D (2016). "Tidytext: Text Mining and Analysis Using Tidy Data Principles in R." *JOSS*, 1(3). doi: 10.21105/joss.00037 (URL:<https://doi.org/10.21105/joss.00037>),