

Using PseudoGravity to Attract People

An Automated Approach to Engaging a Target Audience using Twitter

Soo Ling Lim^{*†} and Peter J. Bentley^{*†}

^{*}Braintree Ltd., London, UK

[†]Department of Computer Science, UCL, London, UK

{s.lim, p.bentley}@cs.ucl.ac.uk

Abstract—We introduce the PseudoGravity tool, an automated social media system that establishes a social media presence in the area of interest of a target audience, identifies target users that are open to connect, engages with them, and elicits a complex response and time investment from them. In this work, we use Twitter as the social media platform and an extensive survey as the activity requiring time investment. We evaluate the tool by using it to find and survey a challenging target – science fiction authors – and compare its results with other methods of automated online surveys. In 28 months, the Twitter account managed by the tool achieved more than 12,000 followers, and achieved monthly Tweet Impressions of more than 250,000. The tool also achieved a high survey response rate of 71% and a completion rate of 83% compared to 30% and 47% achieved by typical online surveys, and high numbers of words and characters entered for questions that required free text input. In addition, this work successfully surveyed more than 500 science fiction writers and gained new understandings of the challenges that e-publishing is bringing to their profession.

Keywords—Social media; Twitter; artificial intelligence; automated engagement; marketing; online survey; e-publishing

I. INTRODUCTION

In today's world of social media and marketing, sales depend increasingly on the ability to identify the correct target audience, and engage with them effectively. Indiscriminate spam marketing, whether by mass-market email, or by messaging in platforms such as Twitter or LinkedIn, is intrusive, unwelcome and ineffective [1]. However, building a well-defined social media profile which is designed to appeal to a specific niche of people and then engaging its audience in a way that does not alienate them, involves considerable time, expertise, and expense – commodities in short supply for the small businesses that need this the most.

The goal of this work is to provide an automated social media tool that enables any given target group of people to be engaged willingly. Instead of misusing artificial intelligence techniques and social media in ill-advised attempts at online manipulation, we wish to provide a service that the target audience desires, and therefore gains by active engagement.

The PseudoGravity tool is our proposed solution. It is an automated social media system that establishes a social media presence in the area of interest of the target audience, identifies target users that are open to connect, builds a relationship with them, and elicits a complex response and time investment from them. Focussing on Twitter, we evaluate the tool on a challenging problem in e-publishing and compare its results

with other methods of acquiring the same response and time investment from people.

II. PROBLEM DOMAIN

Electronic publishing and electronic bookstores are changing the way in which authors reach their audiences. Instead of a few elite publishers who spend considerable sums to publish and market a small number of books, there are now millions of self-published books, which are instantly visible to audiences of hundreds of millions. Instead of authors being an elite few, represented by literary agencies and carefully chosen by commissioning editors, today anyone can write a book and make it available online or via on-demand printing.

This revolution in the publishing industry is causing massive upheavals for everyone. Successful authors must now compete with prolific amateurs with large social media followings. Publishers must cooperate with billion dollar online retailers or see their marketing strategies fail. However, tracking the change of practice and opinion of those within the industry is difficult, because the stakeholders involved are difficult to reach and engage. Such individuals (whether authors or publishers) often receive unsolicited emails or social media messages, and out of necessity will ignore the vast majority. The task is made even more difficult if one wishes to focus on those individuals within smaller niches, for example, authors of science fiction stories or books.

Therefore, in order to investigate the effectiveness of the PseudoGravity tool, we focus on the following challenge: the identification of published authors of science fiction works, and successful engagement with them, with the result that they complete an extensive online questionnaire about their experiences in the publishing industry.

III. BACKGROUND

There are many studies on how organisations use Twitter to engage with their target users and the importance of such engagement. Examples of such organisations include: health agencies [2], non-profits [3], government [4], conferences [5], higher education [6], [7], large enterprises [8], and SMEs [9].

A. Non-Twitter Tools to Find People

Not all approaches that find target audiences make use of existing social media platforms. For example, Lim et al. [10], [11] propose a tool to identify interested parties for large software projects that involve many stakeholders. The tool starts with an initial known set of people and contacts them via email to ask them to recommend new people. It then contacts

the recommended people via email to ask them to make further recommendations, and so on, building a social network with the people as nodes and recommendations as links, and using social network algorithms to rank the people. The approach works well when interested parties are well-connected [12].

B. Classifying User Behaviours on Twitter

If the social network of Twitter is to be useful for identifying target groups, then understanding user interests and intentions through user profiling could be useful. In Penas et al. [13], user interests, intentions, characteristics, behaviours and preferences are inferred from messages using a semantic ontology in which user interests and intentions are characterised. In related work, Wu et al. [14] developed a system designed to generate personalised annotation tags automatically to label Twitter users' interests and concerns, applying TFIDF ranking and TextRank to extract keywords from Twitter messages to tag the user. In Wang et al. [15], the emotional content of tweets was classified, using a combination of unigrams, bigrams, sentiment/emotion-bearing words, and parts-of-speech information.

Kolchyna et al. [16] propose an alternative way to understand user behaviours – through the detection, differentiation and quantification of events in Twitter in order to predict sales. Their study provided empirical evidence that through events differentiation one can clearly identify clusters of Twitter events that contain more information about future sales than the non-clustered Twitter signal.

C. Twitter Tools to Find People

The use of social media to find target groups of people automatically and engage with them is not a common area of study. In Bakshy et al. [17], the attributes and relative influence of 1.6 million Twitter users were studied by tracking 39 million diffusion events that took place on the Twitter follower graph over a two month interval in 2009. Although influencers were discovered, the authors found it difficult to predict which user would have the most future influence and concluded that the use of such influencers could only be harnessed reliably by targeting large numbers of them, thereby capturing average effects.

In a similar vein, Canhoto et al. [18] explored segmentation practices of organisations with a social media presence. The study found that social media are particularly promising in terms of targeting influencers, enabling the cost-effective delivery of personalised messages and engaging with numerous customer segments in a differentiated way.

Messias et al. [19] showed the feasibility of creating Twitter bots that are able to make themselves influential according to established metrics by simple following and tweeting behaviours. While their system did engage with people and resulted in modest success, the indiscriminate behaviours achieved no beneficial effects, and no attempt to

use these methods in order to locate a specific group of users was made.

In summary, while research into social networks is booming, the aims of this work remain novel: the use of social media to accurately target interested parties and engage with them automatically in order to achieve a specific task.

IV. METHOD

The PseudoGravity tool is designed to engage and appeal, never to spam its target audience. Its design follows the recommended social media practices in [20]: *be active, be interesting, be humble, be “charming and personable”, be honest*, and complies with Twitter recommended automation practices¹. It uses the following steps, for the social media platform Twitter:

- 1) An online account is manually created with profile and images that appeal to the target audience.
- 2) A variety of tweetable content is provided, tailored to the interests of the target audience. Social media content is automatically and regularly posted in the area of interest of the target users.
- 3) People are engaged by favouriting selected messages that mention the account name, and retweeting messages that contain specific relevant keywords.
- 4) The followers of similar users are followed. If users that have been followed follow back, then a friendly message is automatically sent designed to encourage a conversation. If they respond and the tool confirms from the interaction that they are members of our target audience, they are sent a message asking them to fill in a questionnaire. If users do not follow back then after a period of time, they are unfollowed to avoid engaging with the wrong people.

For this work, the target audience was science fiction writers who have published books or short stories. The PseudoGravity tool was written in Python using the Twitter API. Fig. 1 shows a block diagram of the platform. We describe each of the four steps in further detail below.

A. Creation of PseudoGravity (Step 1)

To appeal to science fiction authors, a Twitter account was created named *PseudoGravity*. This was given relevant imagery (profile photo and background) and bio text: “I have come to Earth in search of Intelligent Life”, see Fig. 2. (Research shows that profiles with images receive ten times more followers than those without.²)

B. Auto-Generation of Online Content (Step 2)

In the Sharing Component of the tool (Fig. 1) four different sources were used to auto-generate tweets by *PseudoGravity*. Each was designed to appeal specifically to the target audience and provide them with entertainment or news:

¹ <https://support.twitter.com/articles/76915>

² <https://blog.hubspot.com/blog/tabid/6307/bid/5811/Twitter-Accounts-with-a-Profile-Picture-Have-10-Times-More-Followers-Than-Those-Without.aspx>

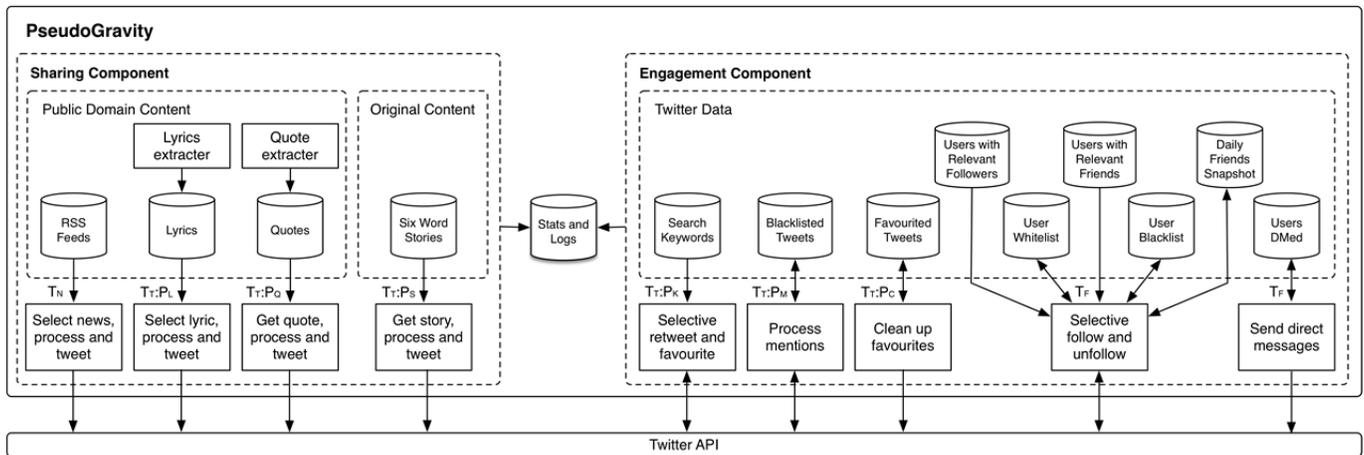


Fig. 1. Block diagram showing the main components of the PseudoGravity tool.

TABLE I. RSS FEEDS USED FOR TWEETS WITH HASHTAGS

RSS Feeds	Hashtag	Handle
http://www.tor.com/rss/frontpage_full	#ScienceFiction	@tordotcom
http://www.forbes.com/entrepreneurs/index.xml	#entrepreneur	@Forbes
http://feeds.feedburner.com/TechCrunch/	#tech	@techcrunch
http://feeds.gawker.com/io9/full	#future	@io9
https://pa.tedcdn.com/feeds/talks.rss	#inspired	@TEDTalks
https://www.wired.com/about/rss-feeds/	#tech	@WIRED
https://www.newscientist.com/rss-feeds/	#science	@newscientist

1) Every $T_N = 2$ hours, RSS feeds from relevant websites such as science magazine New Scientist and science fiction publisher Tor were sampled. If a new item was available the feed title was converted into a character string with appropriate hashtags and Twitter handles added, and URL of the news item (shortened by bitly³). If the resulting string was more than 140 characters, the feed title was shortened. Then it was tweeted. See Table I for details of source sites and corresponding hashtags.

2) Before starting the system, a database of science quotes was generated using BeautifulSoup⁴. These were extracted from a public domain online resource⁵ and parsed using existing HTML tags and regular expression matching to select tweetable quotes based on the following criteria:

- a) Offensive quotes (containing obscenities or foul language) were removed.
- b) Non-English quotes were removed.
- c) Only the last name of quote's author was kept (to meet the 140-character constraint).

d) Corrupted text was removed and formatting is made consistent.

e) The hashtags '#science #quotes' was appended to the end.

f) Any resulting quotes with length > 140 characters were removed.

For example, the quote:

"If we knew what it was we were doing, it would not be called research, would it?" — Albert Einstein

Became the tweet:

"If we knew what it was we were doing, it would not be called research, would it?"-Einstein #science #quotes

Using this database of quotes, every T_T minutes (where T_T is a random number between 90 and 110), with probability P_Q (Table II), the next quote was selected and tweeted. When the end of the database was reached, the system started from the beginning again.

TABLE II. PROBABILITIES FOR USAGE OF DIFFERENT CONTENT

Content Type	Probability Name	Probability
Quote	P_Q	0.185
Six Word Short Story	P_S	0.185
TMBG lyric	P_L	0.185
Retweet of keyword	P_K	0.28
Retweet of mention	P_M	0.15
Clean up	P_C	0.015

³ https://bitly.com/

⁴ BeautifulSoup is a Python library for pulling data out of HTML and XML files: https://www.crummy.com/software/BeautifulSoup/bs4/doc/

⁵ https://www.goodreads.com/quotes/tag/science



Fig. 2. *PseudoGravity* account on Twitter, with username and imagery designed to appeal to the target audience.

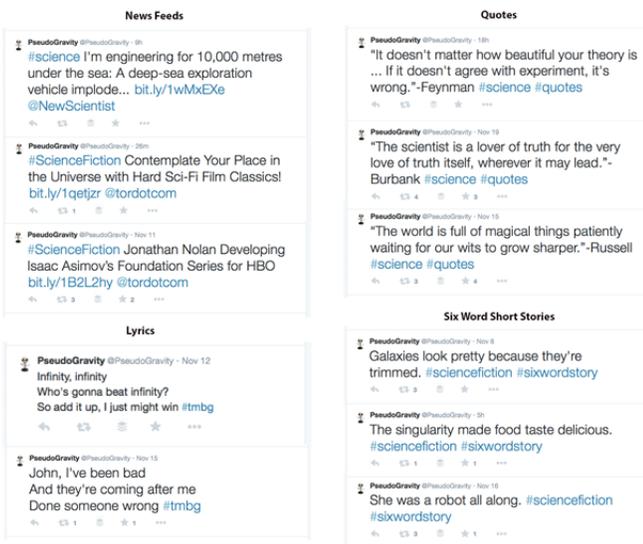


Fig. 3. Examples of the four types of automated tweets by *PseudoGravity*.

3) Before starting, a database of lyrics from the band They Might Be Giants – a band that frequently uses unusual and relevant words in its songs – is automatically generated (one file per track, making 498 files)⁶.

Every T_T minutes, with probability P_L :

a) A random lyric file was opened, a random stanza of lyrics was copied from it, and then (if needed) trimmed down into lines $\leq \text{maxLen}$, where maxLen is a random number chosen from the set $\{120, 120, 120, 120, 100, 100, 100, 80, 80, 40\}$, and where sentences may be trimmed from the back or the front of the list randomly, in order to add variety and increase interest.

b) If a random chunk of lyric was not found that met the maxLen criteria (i.e., the single remaining line is still too long even after trimming), the system tried again up to 10 times.

⁶ Source of lyrics: The TMBG (They Might Be Giants) Knowledge Base <http://www.tmbw.net>. This part of the tool is based on <https://artandlogic.com/2014/01/this-might-be-a-twitterbot/> and uses the code adapted from <https://github.com/bgporter/tmbotg>.

c) If the trimming was successful, the hashtag ‘#tmbg’ was appended to the string, and the result was tweeted.

d) A copy of the lyric that was tweeted (album, track, lyric) was saved to disk.

For example, the stanza:

*Here's hoping you don't (hoping you don't)
Become a robot (clang clang)
Clang clang whoops, too late
Here's hoping you don't (hoping you don't)
Become a robot (clang clang)
Clang clang whoops, too late*

Became the tweet:

*Here's hoping you don't (hoping you don't)
Become a robot (clang clang) #tmbg*

4) Before starting, several hundred six-word stories relating to science fiction were created. These were handwritten by the second author for this work, but could be auto-generated using simple heuristics and keywords. Given that the target audience was science fiction writers, they were designed to be an interesting and relevant form of content. This format was chosen because it is highly relevant for the target audience. For decades, writers have been inspired by Hemingway’s famous six-word tale, “For sale: baby shoes, never worn,” and social media has encouraged new writers to try their own⁷. Some examples of the science fiction six word short stories are:

- *Friends don't sell your brain. Apparently.*
- *Terraformed worlds usually smell like bleach.*
- *Free artificial retinas! (Discreet peripheral advertising.)*
- *Snakes first evolved inside spacecraft pipes.*
- *Bird. Definition: Extinct bipedal phosphate refiner.*
- *43 was actually an error code.*
- *Electricity pylons are skeletons of aliens.*

Every T_T minutes, with probability P_S , the next six word short story was tweeted, with the hashtags ‘#sciencefiction’ and ‘#sixwordstory’ appended. When all stories been tweeted, the system started from the beginning again.

Fig. 3 shows examples of all four types of content automatically posted by *PseudoGravity*.

C. Audience Engagement (Step 3)

Every T_T minutes, with probability P_K , the Engagement Component of the *PseudoGravity* tool (Fig. 1) searched Twitter for tweets that contained the keyword ‘sciencefiction’ or ‘scifi’. If the tweet was not made from the *PseudoGravity* account, was in English, was not offensive (did not contain any obscenities) and if it had fewer than 6 hashtags (to avoid spam), the message was retweeted and the next tweet returned

⁷ http://www.huffingtonpost.com/2014/05/16/six-word-story_n_5332833.html

by the keyword match was favoured. Otherwise the tweet was favoured.

Next, the PseudoGravity tool checked to see if there had been any messages that mentioned the account (i.e., that contained the substring *@PseudoGravity*). For each mention:

a) If the message contained negative keywords, it was added to a blacklist (to prevent favouriting or retweeting negative tweets).

b) If message was not in English, it was offensive, it was blacklisted, it was already retweeted, it was already favoured, or it originated from this account, then it was not retweeted or favoured.

c) Otherwise, if the person who tweeted the tweet was not a friend, or if the words of the message comprised 50% or more user mentions (so it had a lower chance of being relevant) then it was favoured. A copy of the favoured tweet was stored.

d) Otherwise the tweet was retweeted. A copy of the retweeted message was stored.

Finally, with probability P_C , a random number of tweets were selected from the favoured list stored in the system and unfavoured (and removed from the list). After each Twitter API action, the system slept between 2 and 8 seconds in order to simulate human behaviour.

D. Social Networking and Audience Participation (Step 4)

Every T_F minutes (where T_F is randomly chosen between 90 to 150), the Engagement Component of the PseudoGravity tool (Fig. 1) fetched a list of followers (accounts following *@PseudoGravity*) and friends (accounts that *@PseudoGravity* follows). It followed back all followers (except blacklisted accounts), and unfollowed anyone who had unfollowed the account (except whitelisted accounts). It also unfollowed any friends who did not follow back after 5 days. The ‘following behaviour’ of the bot is an example of being personable and honest: reciprocating a follow from a good user is a nice thing to do.

Next the PseudoGravity tool copied the followers of similar accounts (‘ScienceFiction’, ‘tordotcom’, ‘DailySF’, ‘SciFi’, ‘ScienceFicNews’) except those in the blacklist. It also copied the friends of similar accounts. Any not already followed were then followed, subject to the constraint that (friends - followers) < 280. This is to ensure that the system did not follow a disproportionate number of accounts compared to the number that follow back.

A maximum number of 20 new accounts were followed and unfollowed every 5 to 25 seconds (chosen randomly) to simulate human behaviour.

The whitelist was updated by adding names of accounts that we always want to follow, regardless of whether they follow us back, e.g., accounts with large followings that tweeted messages containing relevant keywords. The blacklist was updated by adding names of accounts that displayed bad

behaviour, e.g., their account was suspended by Twitter, they were no longer active, they sent an inappropriate message containing obscenities, or spammed (too many messages containing too many hashtags and/or account mentions).

The final activity by the PseudoGravity tool was to encourage audience participation from the members of its growing social network. This was achieved by sending a friendly direct message to any user that is both a friend and a follower of *@PseudoGravity*:

“Thanks for the follow. Do you like reading science fiction?”

Table III shows the subsequent messages that were sent, depending on the response. Each message was sent at most once to the same user. It should be noted that the main aim of the system was to find science fiction writers who will click on the survey link and participate in the survey, but importantly the system never tweeted the survey link publicly. The link was only provided by private direct message after the participants explicitly said they would be happy to help. This gentle introduction to the task follows known psychology: it is better to have the user invest a small amount of time in a conversation before asking for a larger investment of time doing a survey [21].

If the user successfully obtains the link, follows it, and completes the survey – which can take considerable time for each person – this is a successful engagement with the target audience.

TABLE III. DIRECT MESSAGE RESPONSES

Reply Contains	PseudoGravity Response
Write science fiction	We are running a survey about science fiction writers and the problems they face. Would you like to participate?
Recommend books?	I recently read Zones of Thought by Vernor Vinge - quite enjoyed it.
Favourite genre?	I am a sucker for space opera.
What is your name?	Where I am from, we don't have names.
Are you bot / alien?	I do not know what a bot is. I am not from your world. Does that make me an alien?
Are you artificial intelligence?	Thank you for your assessment of my intelligence. I am not any form of automaton. I am on a mission to find intelligent life.
How is the search going? [in response to the profile bio]	It's exhausting.
Have you published?	I have published some short stories a while ago.
Happy to help with survey	Great! Thanks. The link to the survey is ...
Do I qualify?	We are targeting all types of authors, so your input would be very useful.
Does my work qualify?	We target science fiction in the widest sense, including fantasy, horror, speculative, slipstream and others.
Will look but busy	We really appreciate any time you can spend. The link to the survey is ...

V. SURVEY

In addition to investigating the utility of an automated approach to finding a target audience using social media, the secondary purpose of this research was to conduct a survey of the target audience: science fiction writers. The aim of the survey was to discover the main factors affecting writers of science fiction as e-publishing becomes more prevalent. Do science fiction authors find it easier or harder to have their work published? Is e-publishing helpful or is it hindering their success?

The questionnaire contains 27 questions, with two that permit the entry of free text, and one optional question asking for email details of the participant. The full questionnaire is provided in Appendix A.

VI. EXPERIMENT

A. Setup

The *PseudoGravity* Twitter account was created on 27 May 2014 and the *PseudoGravity* tool was activated on the same day. The *PseudoGravity* tool was deactivated on 30 Aug 2016.

B. Results

Figs. 4 to 6 show the results from the *PseudoGravity* Twitter account. Fig. 4 shows the steady increase of followers that the account gained. By the end of the experiment, *PseudoGravity* had more than 12,000 followers. This is a good total; for example, the Twitter account for the British Computer Society and Chartered Institute for IT @bcs had 17,000 followers at the time of writing. The *PseudoGravity* account saw a consistent rate of follower growth; for example, in the year 2015 growth was around 66% (an increase of 5616 to 9325), which also compares well with Twitter as a whole, which had follower growth in the year of 2015 of 23%⁸.

Fig. 5 shows the Twitter impressions for the tweets produced by the *PseudoGravity* tool. Monthly impression totals were consistently around 150,000 and peaked at more than 250,000. This is a significant number – in 2015, the average cost per 1000 impressions (CPM)⁹ for Twitter was reported to be \$12.16. So to achieve the same tweet impressions would have cost around \$3,000 USD per month.

Fig. 6 shows the cumulative number of surveys started and completed over time. The average duration spent by participants completing the survey was 29.6 minutes.

C. Analysis

Ultimately the purpose of the *PseudoGravity* account was to maximise the number and quality of survey responses. One possible method for evaluating the results achieved by this approach would be to compare our survey responses with the responses collected by a professional survey research company. We contacted Cint¹⁰, an established ISO certified panels company for conducting opinion and social research and was informed that they would be unable to provide any results as authors were not profiled on their system.

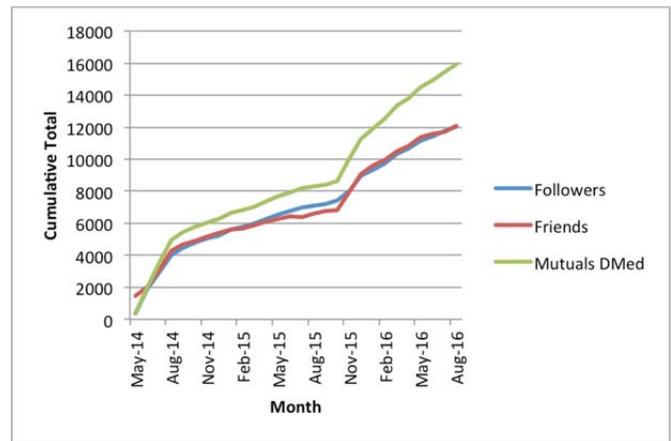


Fig. 4. Increase over time of followers (accounts following *PseudoGravity*), friends (accounts that *PseudoGravity* follows) and mutuals (mutual friends and followers) that are sent direct messages by *PseudoGravity*. The reduced rate from August 2014 until October 2015 was caused by decreasing the number of accounts that the system could follow and unfollow at a time to 10, down from 20, in order to satisfy a temporary change in the Twitter API.

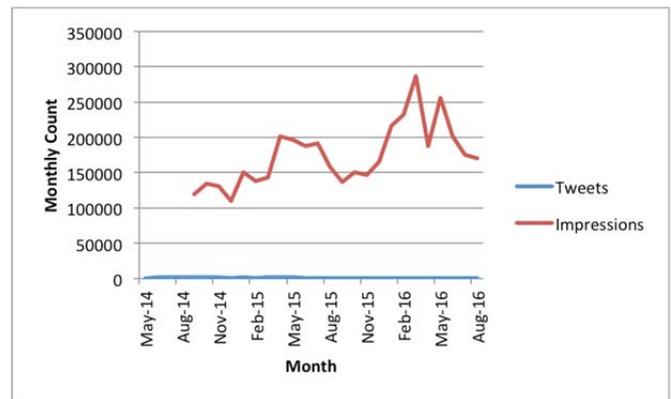


Fig. 5. Monthly tweets and impressions (the number of times each tweet was seen) over time. On average, 1244 messages were tweeted per month, and each tweet has an average of 149 impressions. Data for impressions was only made available by Twitter from September 2014.

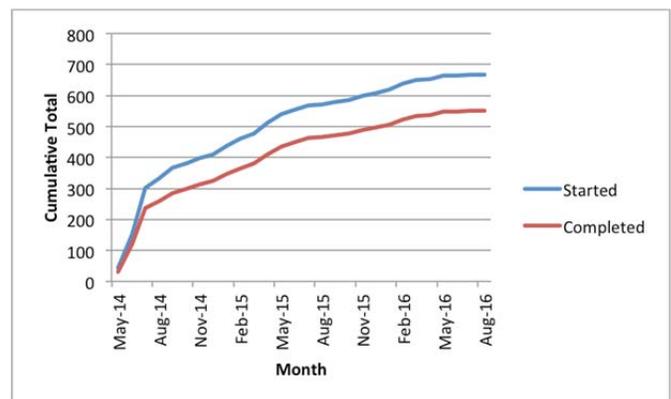


Fig. 6. Cumulative number of surveys started and completed over time.

⁸ <http://pages.trackmaven.com/rs/251-LXF-778/images/trackmaven-social-media-inflation-infographic.pdf>

⁹ <https://www.ft.com/content/91a471be-ea87-11e4-96ec-00144feab7de>

¹⁰ <http://cint.com/>

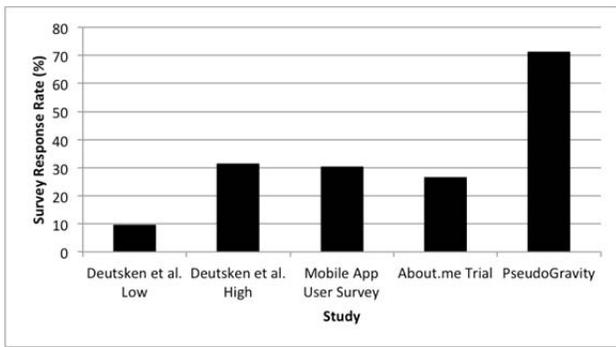


Fig. 7. Comparison of survey response rates reported in Deutsken et al. [22], mobile app user survey [23], and the rate achieved by the PseudoGravity tool.

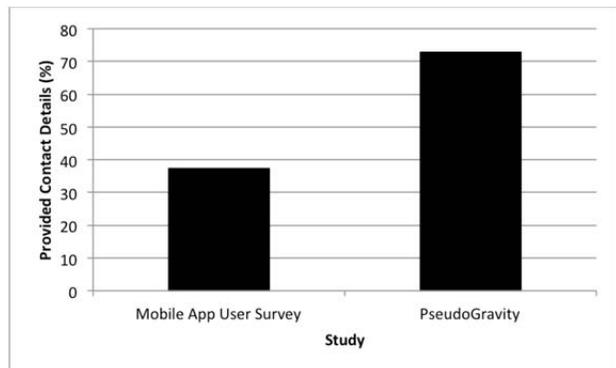


Fig. 8. Comparison of percentage respondents who provide contact details reported in the mobile app user survey [23] and the rate achieved by the PseudoGravity tool.

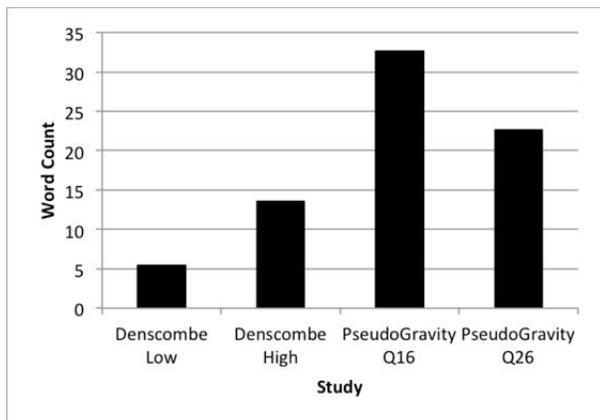


Fig. 9. Comparison of average word count for free text entry by participants reported in Denscombe [24] and the number achieved by the PseudoGravity tool for survey questions 16 and 26.

However, it is possible to compare the effectiveness of this approach by comparing with other online surveying methods. In Deutskens et al.'s study of the response rate of online surveys with different configurations (e.g., short vs. long, donation to charity vs. lottery incentive, early vs. late reminder), they found that the response rate ranged from 9.4% to 31.4% [22]. Similarly, in a recent survey of more than 10,000 mobile app users [23] that was conducted using Cint panel, the response rate was 30.2%.

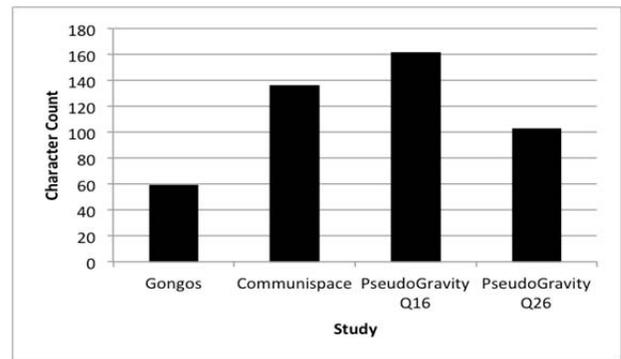


Fig. 10. Comparison of average character count for free text entry by participants reported in the literature [25] and the number achieved by the PseudoGravity tool for survey questions 16 and 26.

To provide a more direct comparison with this approach, between 26 June 2014 and 2 July 2014, we used keyword search [science fiction, author] and [science fiction, writer] on social network About.me and contacted 204 people whose profile indicated they were science fiction authors. Of these, 54 agreed to complete the survey, giving a response rate of 26.47%. In contrast, for the PseudoGravity tool, 769 responded to the auto direct message, were deemed science fiction authors, and so were asked if they wanted to do the survey. Out of those, 547 went on to do the survey, giving an impressive response rate of 71%¹¹, see Fig. 7.

In the mobile app user survey which used Cint panel [23], 37.4% of participants volunteered their contact details. For this approach, 72.8% volunteered their contact details (Fig. 8). In the same survey, the percentage of completed surveys was 79.2% (each participant receives payment on completion), but only 47% provided useable results – the other completed surveys were filled with unusable responses (e.g., “N/A” or garbage) as the respondents were unsuitable but still wanted to complete the survey. In contrast, 82.6% completed the survey for the PseudoGravity tool with all answers usable.

Likewise, the amount of effort spent completing the surveys was also high for this approach. In [24] the average word count for open-ended questions (where respondents may write as much or as little free text as they chose) ranged on average from 5.5 to 13.6 words. In this study, there were two open-ended questions; the average word count for the first was 32.7 words, and for the second was 22.7 (see Fig. 9). The maximum number of words provided as a response was 471 words (about half a page of text). In another study [25] the average character count for free text in surveys was 136 (excluding whitespace) for existing members of an online community. In our approach, the average character count (excluding whitespace) was 161 and 102.6 (see Fig. 10).

These results all clearly show the advantage of using social media to target the correct audience, and provide that audience with useful content to ensure they engage and participate.

The PseudoGravity tool ran for 28 months on the *PseudoGravity* account and proved to be a highly effective method for automatically engaging a target audience. These

¹¹ This figure does not include surveys performed as a result of the link being provided via an alternative source, e.g., in the About.me experiment.

results were further corroborated by the use of the PseudoGravity tool for three projects in various domains: the identification and engagement of activity organisers for an activity marketplace, budding online musicians for a music startup, and gift card users for small businesses. In all cases, the tool built social media accounts of several thousand followers and obtained contact details of many potential customers. In one instance, the tool even managed to win a prize for the account (a children’s board game). The PseudoGravity tool was also used on the first author’s own Twitter account, with the result that a social media analytics company rated the Twitter account as the top 20 safety and technology influencers in the British automotive industry, alongside people from Financial Times, Forbes, and New York Times [26]. Their ranking methodology takes into account resonance (how much engagement they are driving), relevance (how many times they tweet on topic), reach (how many followers they have) and reference (how many twitter lists they appear on) [26].

One final method for assessing the performance of the system is via the reaction of the audience themselves. Although the system contains a blacklist designed to find complaints tweeted at the PseudoGravity tool so that they are not favoured or retweeted, there were no complaints ever sent. The only negative message ever received was:

@PseudoGravity I notice you send direct messages but cannot receive them. Netiquette says: to give is to receive... Yes, I read #scifi.

This on analysis was caused by a temporary glitch in Twitter that disallowed him from sending a direct message to the *PseudoGravity* account and the misunderstanding was quickly resolved. (The same user still follows and retweets messages from *PseudoGravity* to this day.)

The lack of complaints, and indeed, the extremely positive reaction to the social media accounts, reinforces the central theme of this work: it is possible to use an automated tool as an intelligent curation device to provide positive and beneficial content for a target audience. This is not a spamming bot. This is a method that finds like-minded people and provides them with content that they enjoy.

VII. SURVEY RESULTS

The main focus of this work has been to describe the PseudoGravity tool and provide evidence of its effectiveness in accurately targeting interested parties. In this section, we briefly summarise the results of the survey.

A total of 668 responses were received. Respondents were located in 35 countries across the world (Fig. 11), and ranged from new authors to multi-time Hugo award winner with over 20 novels in print to successful self-published author averaging \$240k per year. Of the 522 respondents who completed the survey, 92% wrote science fiction, 64% also wrote fantasy, and 28% horror. Nearly 20% had published between 3 and 7 books or novellas and 7% had published more than 7.

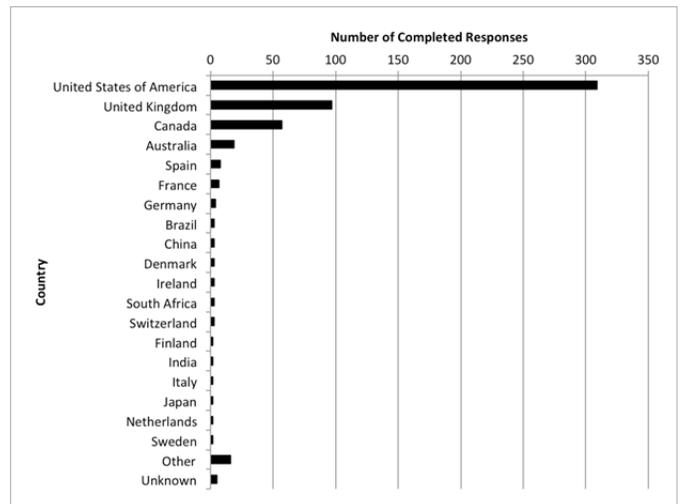


Fig. 11. Distribution of respondents who completed the survey (N=552) by country. Total number of countries is 35. Countries in the “Other” category are Belgium, Bosnia and Herzegovina, Czech Republic, Greece, Isle of Man, Israel, Malta, Mexico, New Zealand, Norway, Philippines, Portugal, Russia, South Korea, Trinidad and Tobago, and United Arab Emirates.

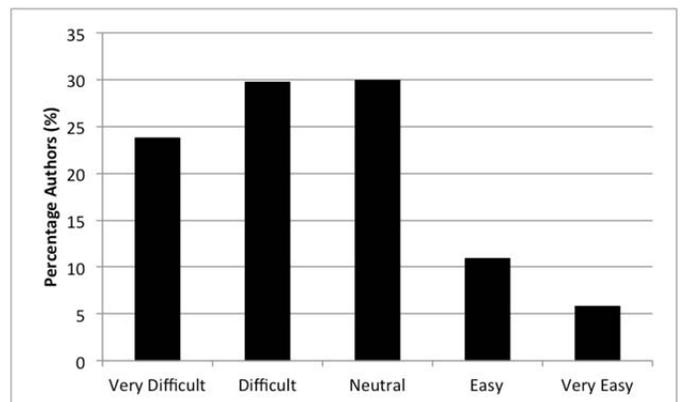


Fig. 12. Responses to Question 10 of the science fiction survey: How easy do you find it to have your story published in your chosen market?

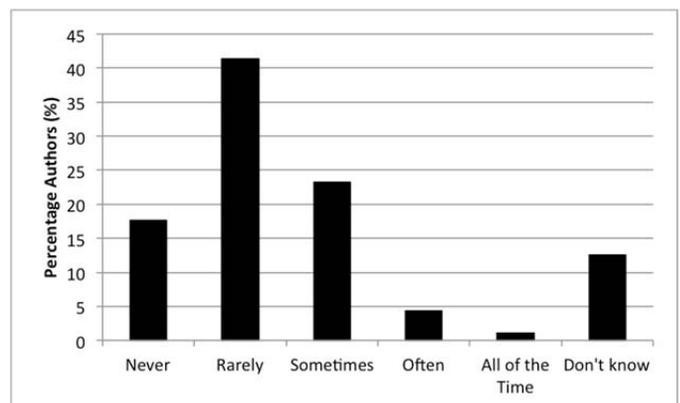


Fig. 13. Responses to Question 14 of the science fiction survey: How frequently do you reach most of your potential audience?

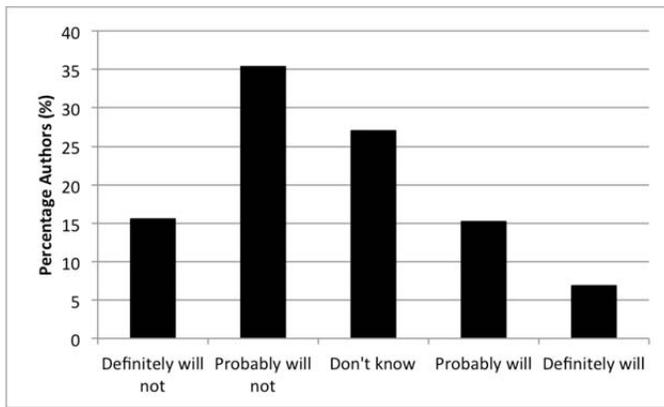


Fig. 14. Responses to Question 15 of the science fiction survey: Over the next ten years, do you believe you will be able to make a living from writing science fiction alone?

Despite the new opportunities in e-publishing, most authors felt it was not easy to publish their work in their chosen market, Fig. 12. Despite the hugely increased opportunities that now exist for authors to reach their readers via social media or e-bookstores, most authors felt that they were not reaching most of their potential audience, Fig. 13, and believed that the increased competition caused by self-publishing made it ever harder to stand out from the crowd. When asked whether they would be able to make a living from writing science fiction in the next decade, most were pessimistic, Fig. 14.

In summary, the survey appears to indicate that e-publishing is not improving the lives of writers and their ability to make a living from their trade. One of the frequently reported problems was oversaturation in the market. It is clear that new innovations are needed to address these issues. The results of the survey have been released to the science fiction author community.

VIII. CONCLUSIONS

Social media offers a new channel for providing awareness of services and products to people. There are many dirty tricks using social media being played in the name of 'advertising': spam, posting misleading 'facts', answering questions posed to competitors, even creating fake accounts to artificially boost apparent social media followings.

In this work, we described an alternative approach. The PseudoGravity tool is an automated social media system that establishes a social media presence in the area of interest of the target users, identifies target users that are open to connect, builds a relationship with them, and elicits a complex response and time investment from them. It achieves this not by fakery or artifice, but by providing new, useful and relevant content to the target users, and automatically refining its social network to ensure only those with genuinely shared interests are retained.

Focussing on Twitter, we evaluated the tool on a challenging problem in e-publishing and compared its results with other methods of acquiring the same response and time investment from people. The results were impressive, with the Twitter account achieving more than 12,000 followers, and the tweets achieving high numbers of Tweet Impressions. When analysing the results of the engagement in terms of surveys

completed, the tool achieved a remarkable response rate of 71%, a completion rate of 83%, and higher than average numbers of words and characters entered for questions that required free text input. This success was also duplicated several times over when the same tool was used for three separate projects in different domains.

Finally, in addition to demonstrating a highly successful approach for accurately targeting interested parties and engaging with them automatically, this work successfully surveyed more than 500 science fiction writers and gained new understandings of the challenges that e-publishing is bringing to their profession.

REFERENCES

- [1] G. Stringhini, C. Kruegel, and G. Vigna, "Detecting spammers on social networks," in Proc. of the 26th Annual Computer Security Applications Conference (ACSAC), 2010, pp. 1-9.
- [2] S. Bhattacharya, P. Srinivasan, and P. Polgreen, "Engagement with health agencies on twitter," PLoS One, vol. 9, e112235, 2014.
- [3] K. Lovejoy and G. D. Saxton, "Information, community, and action: how nonprofit organizations use social media," Journal of Computer-Mediated Communication, vol. 17, pp. 337-353, 2012.
- [4] F. D. L. Wigand, "Twitter in government: Building relationships one tweet at a time," in Proc. of the 7th International Conference on Information Technology: New Generations (ITNG), 2010, pp. 563-567.
- [5] W. Reinhardt, M. Ebner, G. Beham, and C. Costa, "How people are using Twitter during conferences," Proc. of the 5th EduMedia Creativity and Innovation Competencies on the Web, pp. 145-156, 2009.
- [6] R. Junco, G. Heiberger, and E. Loken, "The effect of Twitter on college student engagement and grades," Journal of Computer Assisted Learning, vol. 27, pp. 119-132, 2011.
- [7] E. Kassens-Noor, "Twitter as a teaching practice to enhance active and informal learning in higher education: The case of sustainable tweets," Active Learning in Higher Education, vol. 13, pp. 9-21, 2012.
- [8] S. Rybalko and T. Seltzer, "Dialogic communication in 140 characters or less: how Fortune 500 companies engage stakeholders using Twitter," Public Relations Review, vol. 36, pp. 336-341, 2010.
- [9] M. Bulearca and S. Bulearca, "Twitter: a viable marketing tool for SMEs?," Global Business and Management Research, vol. 2, pp. 296-309, 2010.
- [10] S. L. Lim, D. Quercia, and A. Finkelstein, "StakeSource: harnessing the power of crowdsourcing and social networks in stakeholder analysis," in Proc. of the 32nd ACM/IEEE International Conference on Software Engineering-Volume 2, 2010, pp. 239-242.
- [11] S. L. Lim, D. Damian, and A. Finkelstein, "StakeSource2.0: using social networks of stakeholders to identify and prioritise requirements," in Proc. of the 33rd International Conference on Software Engineering (ICSE), 2011, pp. 1022-1024.
- [12] S. L. Lim, D. Damian, F. Ishikawa, and A. Finkelstein, "Using Web 2.0 for stakeholder analysis: StakeSource and its application in ten industrial projects," in Managing Requirements Knowledge: Springer, 2013, pp. 221-242.
- [13] P. Penas, R. Del Hoyo, J. Veja-Murguia, C. Gonzalez, and S. Mayo, "Collective knowledge ontology user profiling for Twitter--automatic user profiling," in Proc. of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2013, pp. 439-444.
- [14] W. Wu, B. Zhang, and M. Ostendorf, "Automatic generation of personalized annotation tags for twitter users," in Proc. of the Human Language Technologies: 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics, 2010, pp. 689-692.
- [15] W. Wang, L. Chen, K. Thirunarayan, and A. P. Sheth, "Harnessing twitter" big data" for automatic emotion identification," in Proc. of the 2012 International Conference on Privacy, Security, Risk and Trust

- (PASSAT) and 2012 International Conference on Social Computing (SocialCom), 2012, pp. 587-592.
- [16] O. Kolchyna, T. T. Souza, P. C. Treleaven, and T. Aste, "A framework for Twitter events detection, differentiation and its application for retail brands," in Proc. of the IEEE Future Technologies Conference (FTC), 2016, pp. 323-331.
- [17] E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts, "Everyone's an influencer: quantifying influence on twitter," in Proc. of the 4th ACM International Conference on Web Search and Data Mining (WSDM), 2011, pp. 65-74.
- [18] A. I. Canhoto, M. Clark, and P. Fennemore, "Emerging segmentation practices in the age of the social customer," *Journal of Strategic Marketing*, vol. 21, pp. 413-428, 2013.
- [19] J. Messias, L. Schmidt, R. Oliveira, and F. Benevenuto, "You followed my bot! Transforming robots into influential users in Twitter," *First Monday*, vol. 18, 2013.
- [20] A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of social media," *Business Horizons*, vol. 53, pp. 59-68, 2010.
- [21] R. B. Cialdini, *Influence: The Psychology of Persuasion*. HarperBusiness, 2007.
- [22] E. Deutskens, K. De Ruyter, M. Wetzels, and P. Oosterveld, "Response rate and response quality of internet-based surveys: an experimental study," *Marketing Letters*, vol. 15, pp. 21-36, 2004.
- [23] S. L. Lim, P. J. Bentley, N. Kanakam, F. Ishikawa, and S. Honiden, "Investigating country differences in mobile app user behavior and challenges for software engineering," *IEEE Transactions on Software Engineering (TSE)*, vol. 41, pp. 40-64, 2015.
- [24] M. Denscombe, "The length of responses to open-ended questions: a comparison of online and paper questionnaires in terms of a mode effect," *Social Science Computer Review*, vol. 26, pp. 359-368, 2008.
- [25] R. Mackenzie and J. W. Schlack, "Character Counts: A Comparison of Mobile and Online Open-Ended Survey Responses," *Communispace Corporation*, 2012.
- [26] Onalytica, "The British Automotive Industry: Top Influencers and Brands," [Online]. Available: <http://www.onalytica.com/wp-content/uploads/2016/08/British-Automotive-Industry-Top-Influencers-Brands.pdf>, 2016.

APPENDIX A

Science Fiction Survey: Q16 and Q26 were open-ended questions. All other questions had between 3 and 8 options to choose from, with many also allowing the choice of "other" and entry of free text. For brevity, only the questions are shown here.

- Q1. Generally in which genres do you write? Choose as many answers as you like.
- Q2. Approximately how many short stories have you had published?
- Q3. Approximately how many books or novellas have you had published?
- Q4. Approximately how many of your stories (of any length) have not yet been published?
- Q5. In addition to writing fiction, do you make a living from any of the following activities? Choose as many answers as you like.
- Q6. At which of the following computer skills are you proficient? Choose as many answers as you like.
- Q7. How much do you use online websites (blogs, Twitter, Facebook, etc) to market yourself?
- Q8. How easy is it for you to think of new ideas for stories?
- Q9. Before submitting, how easy it is for you to obtain useful feedback from readers that helps you improve your work?
- Q10. How easy do you find it to have your story published in your chosen market?
- Q11. How satisfied are you with the response time after submission for your chosen market(s)?

- Q12. How satisfied are you with the amount of money you receive for your work?
- Q13. How satisfied are you with your recognition and visibility as an author?
- Q14. Compare the total potential audience for your work with the number of people who read your work. How frequently do you reach most of your potential audience?
- Q15. Over the next ten years, do you believe you will be able to make a living from writing science fiction alone?
- Q16. What do you think is the biggest problem facing science fiction writers today?
- Q17. How useful would it be to have a way to connect with other writers to help create and refine new ideas for stories?
- Q18. How useful would it be to have a way to connect with your readers to help create and refine new ideas for stories?
- Q19. How useful would it be to connect with other writers to obtain feedback about a new story before you submit?
- Q20. How useful would it be to connect with your readers to obtain feedback about a new story before you submit?
- Q21. In an ideal world, how long should you have to wait for the response after submission to a market?
- Q22. There are many kinds of market available for your writing. How acceptable do you find these different types?
- Vanity press (you pay to publish, no royalties)
 - Famous name magazine, but terrible payment
 - Self-publishing (e.g. Kindle, iBooks)
 - Online mag with average payment + profit sharing of print anthologies
 - Indy publisher, with 50:50 profit sharing
 - Online mag with crowd-sourced payments (if story is great, so is payment)
 - New magazine with no reputation, but great payment
 - Online mag with reading fee, but great monetary prize if published
- Q23. Think about how many readers actually read each of your stories. Imagine multiplying that number by 1000. To achieve this goal, how acceptable do you find each of the following options?
- Pay a social media expert to improve my online presence.
 - Pay to attend more events and network like crazy.
 - Try to find a good literary agent.
 - Spend a lot of my time becoming a blogging and Tweeting demon online.
 - Write and submit many stories in the hope of being discovered.
 - Try to improve my own online presence using free Websites.
 - Put some of my stories online for free to attract readers.
 - Pay for software which improves my social networking presence and communications.
 - Try to connect with many reputable authors and editors online.
- Q24. Think about how many words you write each year. If you wanted to make a living from writing science fiction, based on your current output, how much do you need to be paid per word?
- Q25. Who should pay the marketing and social media costs in order to increase the visibility and reputation of authors? Choose as many answers as you like.
- Q26. Have you achieved all your writing ambitions? If not, what is the biggest hurdle in your way?
- Q27. If you are interested in this project or want to know the results of the survey, please provide your email address.