

Motivating Online Publication of Scholarly Research Through Social Networking Tools

Vincent S. Smith^{1*}, Daphne Duin², David Self¹, Irina Brake¹ & Dave Roberts¹

¹The Natural History Museum, Cromwell Road, London, SW7 5BD, U.K.

*vince@vsmith.info;

²Museum National d'Histoire Naturelle, CP 43 (Dir. des Collections), 57, rue Cuvier, 75231 Paris cedex 05, France

Abstract. We conducted interviews amongst the primary users of a social networking tool used by natural history scientists to share and manage information on biodiversity. Our goal was to understand the impact, motivations and barriers experienced by respondents. Results indicate that users are motivated to engage through the social and technical innovations offered to enhance collaboration. However, users struggled to engage and sustain these collaborations. In these instances communication of research became a primary motivator. In effect this system is being used by most respondents as a digital surrogate for paper publications, although a substantial minority do collaborate online. The primary barrier of use was the users' lack of time to engage with the system. We infer that this is due to a lack of institutional or career incentives. Motivating sustained use would be most effective if the traditional principles of scientific publications could be integrated into the system.

1 Introduction

Recent years have seen the emergence of new ideas about the publication of scholarly research with talk of a 'crisis' in publishing and weaknesses in the peer-review system. As Open Access publishing becomes ever more widespread [1], [2] even more radical ideas for the 'opening' of scholarly communication are being proposed. One outcome is the notion of 'Open Science' [3] with its advocacy of more open scientific knowledge production and publishing processes [4], [5], inspired by discourses developed in 'Free/Open Source Software' and 'Creative Commons' movements [6], [7], [8]. Web 2.0 is widely seen as providing the technical platform essential to this 're-evolution' of Science [9], [10].

As an experiment in Open Science a team of scientists from the European Distributed Institute of Taxonomy (EDIT – <http://www.e-taxonomy.eu/>) have developed a Web 2.0 Virtual Research Environment (Scratchpads – <http://scratchpads.eu>) that enables communities to collaborate in the production of websites supporting natural history science. These cater to the particular needs of multiple research communities through a common database and system architecture [11]. Registrants assume responsibility for the contents of each site, which (on approval) are instantiated at web domains of their choice. This tool facilitates the rapid construction, curation and publication of content rich web pages about any

taxonomic group. The framework currently serves more than 1,800 registered users across more than 160 sites spanning scientific, amateur and citizen science audiences. Sites range in function from supporting the work of societies and conservation efforts to the production and dissemination of taxonomic checklists, peer reviewed journal articles and electronic books.

Studies of biological taxonomy, systematics and biodiversity (herein referred to as 'natural history') are particularly suited to new forms of scholarly communication. The discipline has made considerable use of Information and Communication Technologies (ICT) in recent decades and its community of scientists is increasingly dependent on these systems. Databases for cataloguing specimen collections with their derivative data and computerized methods for inferring the evolutionary relationships of organisms are central to most aspects of natural history research. Databases have become a significant means of communicating taxonomic information, and latterly there exist many online taxonomic information services including online identification guides, natural history observation catalogues and nomenclatural resources. Dwindling financial support for natural history research has also been used by funding agencies to drive this transition, creating a motivated (albeit sometimes skeptical) community within which certain individuals are increasingly willing to engage with new forms of scholarly communication.

'Scratchpads' were founded in March 2007 and as part of their evaluation and monitoring processes a survey was commissioned to understand the i) impact, ii) motivation, and iii) barriers to the use of Scratchpads experienced by their user community. As part of this work we also sought to gain insight into the profile of users and evaluate the efficacy of the survey as a mechanism to generate these data. This report focuses on the motivation and barriers of use aspect of this survey which are central to user uptake and adoption by the scholarly community. A central goal of the Scratchpad development team was to build a system that could motivate individual researchers in the generation, management and dissemination of their own data for their own needs, while empowering a wider constituent of potential users who are free to repurpose this information. The survey was conducted to understand whether these objectives have been met, and to identify refinements in the technical and social functionality of the system that might be refined to better achieve the project's objectives.

2 Methods

Candidates for the survey were self-selected from the pool of Scratchpad site "maintainers". These are a category of super user who usually instigate the site and have administrative privileges that enable them to invite other users to join the community. Maintainers set levels of access for other users and take ultimate responsibility for their site content. Each Scratchpad community has its own organizational structure and division of labour, with several sites forming "communities of one". In these circumstances the site maintainer is the sole contributor to a site. Generally Scratchpad maintainers form a subset of Scratchpad

users that engage with a wider range of Scratchpad functions and / or are more connected with the specialist users working within a particular community.

For the survey we used the maintainer contact list of July 2009. This included 107 Scratchpad maintainers caring for 129 sites (several users are maintainers for more than one site). All maintainers were invited to participate in the survey and invitations were sent by email. Interviews were booked using an online calendar that integrated with a Scratchpad, enabling users to quickly arrange a slot using a technical environment with which they were familiar. The survey was also announced on several websites, mailing lists and newsletters used by Scratchpad contributors and members of the European Distributed Institute of Taxonomy. Maintainers were incentivised to engage with the survey with the undertaking that feedback would directly influence the Scratchpad development team in the enhancement of the software.

Over a period of four weeks in September 2009, 46 maintainers responsible for 61 Scratchpad sites were interviewed by a sociologist (face to face and by telephone) applying a combination of 65 open and closed questions (see <http://www.e-taxonomy.eu/scratchpadsurvey>). Respondents included a small number of maintainers that had signed up to a site but later abandoned it (so called “leavers”). Interviews were structured around the following three themes:

- 1) How has this technology impacted on the way users communicate and collaborate with their peers, specifically with respect to the way users organize and publish their “data” in the broadest sense?
- 2) What is the general attitude of Scratchpad users towards new technologies in support of scholarly activities, as demonstrated by their adaptation and use of other virtual tools?
- 3) What are the technical and social barriers for adoption faced by users when working with the Scratchpads and other virtual research tools?

Questions were designed to capture current attitudes and patterns of adoption in addition to identifying problems, needs and aspirations of users. These were initially refined through test interviews conducted with a small set of Scratchpad users back-to-back with a Scratchpad training session in Leiden, Netherlands. Subsequent in-depth, semi-structured interviews were conducted with respondents in order to explore the uses they were making of the system, their experiences, motivations and their perceptions of barriers and drivers to adoption. Closed questions were used where possible, mainly to verify information already given or to test preconceived notions captured from a rolling program of user feedback and engagement with the development team. Open-ended questions were used to investigate user opinions and attitudes for issues on which we had no prior information and to ensure the interview process did not stifle their response. All interviews were recorded and responses were transferred to an online system for management and analysis including transcriptions of key points from open-ended questions.

Because Scratchpads are built, developed and used by members from the taxonomic community, and in particular by staff from Natural History Museum, London (NHML), the interviewer was recruited outside this community to guarantee impartiality and objectivity to data collection. In order to detect possible sampling

bias, the respondents demographic data (age, sex, academic affiliation, country of origin, taxonomic focus of their Scratchpad) was compared with that of the total Scratchpad maintainer population, as determined from statistics captured during their point of sign-up with the system.

3 Results

Interviews had an average duration of 37 minutes (17m minimum, 90m maximum). No significant bias in the sampling was detected, with the notable exception of institutional and country affiliation. The strong user base of the Scratchpad system at NHML meant that a high proportion (26%) of all Scratchpad maintainers are based at this institution. Of these users 19 (41%) were included in the survey. This also created a bias in country affiliation with 31 (67%) of maintainers in the survey based in the UK, against a backdrop of 45 maintainers (42%) present in the overall Scratchpad maintainer population. Other demographic factors showed no significant bias in the survey.

3.1 Impact of Use

The results reveal that the most common reason maintainers registered for a Scratchpad is because they needed a tool for collaborative activities (37%). Additional reasons include the need for a tool to communicate research (24%), a general interest in bioinformatics (20%) and for data management purposes (17%). These data broadly match supplementary questions about how respondents use their Scratchpad. 67% report that they use their Scratchpad to communicate research and 64% see this as the primary benefit of their Scratchpad. 62% also report that their site is used for collaborative activities and see this as a secondary benefit. Additional uses reported include archiving individual data (33%); automatically generating data presentations from data records (21%); and blogging (17%), especially with respect to reporting narratives on fieldwork activities. Some users additionally participate in group blogs (7%), and a significant number (40%) use their Scratchpad to manage and collaborate in the production of shared bibliographic reference lists. Collectively these data counter the traditional perception of taxonomists as loners that are conservative in the research methods [12]. Indeed, some maintainers report using their Scratchpad to specifically learn about other researchers work.

When asked in an open question about spin-off activities that have specifically resulted from users working with their Scratchpad, narrative responses include invitations to give presentations, requests for joint authorship on publications and collaboration in writing grant proposals to seek further funding. 64% of respondents noted that they have yet to experience any spin-off benefits. Nevertheless, a substantial number of users report changes in their working practices. We asked “if the virtual tools used had changed the way respondents worked with others”. Respondents were asked to compare this with their personal working practices before and to give their opinion of this change. The 38 responses to this question range from

users reporting a “slight change”, to collaboration as being “completely different”. Significant changes reported include improvements in communication efficiency; scaling-up communication to reach a larger audience and the possibility to participate in complex communication processes with different groups of people across different document formats. One young respondent stressed that they would not have been involved in this research field without these tools on the Internet.

Overall, evidence for regular and sustained engagement with the Scratchpads remained high. 54% of users reported logging in to one or more of their sites in the past week and a further 20% logged in within the last month. Just 24% have only logged in between 1 and 6 months ago, and a single user did not log-in again after signing up. We attempted to speak with more Scratchpad maintainers that had abandoned their site completely. We found 4 “leavers” (included in the set of 46 respondents) willing to participate in the interview. Three of these four cited the complexity of the system as their primary reason for abandoning the site. Other reasons given include software bugs, limited time and the fact that the software did not conform to user expectations.

3.2 Attitudes to New Communication Technologies

We asked respondents about their attitudes to new technologies supporting the communication of their research, especially with respect to collaborative ways of working. Where appropriate, these questions were followed by supplementary responses or actions providing behavioral evidence of these perceptions. Exactly half the Scratchpads covered in the survey had just a single user (the maintainer) and no other active members of the site. However, 39% reported 2-10 users and two sites reported more than 10 users. These responses were validated by the interviewer examining the public statistics on the specific Scratchpad sites after the interview. Maintainers of single user sites were asked why they were the only active user. The three most common responses relate to concerns over the loss of quality control; that it was premature to invite others because the site was insufficiently developed; and because the maintainer perceived the potential community of contributors to be too small.

Supplementary questions were asked about how often respondents use other virtual research tools, patterns of off-line working, and general mechanisms for communicating their research. Perhaps not surprisingly, a substantial majority (83%) of respondents used other virtual research tools (specifically web based software other than e-mail or Scratchpads) as part of their research activities. Of these users, 87% access these virtual research tools daily and 83% use these to collaborate. These collaborative patterns of working are broadly mirrored by off-line activities. All but one respondent stated that they work off-line with collaborators at other institutions and most (78%) also work as part of local research teams, participating in local research networks.

Within the Scratchpads a majority of collaborations (48%) were with other specialists working on the same taxonomic group. Others were in support of more information driven curation activities focusing on particular categories of data (35%), while a substantial number (30%) were engaged in interdisciplinary virtual research

working with other categories of scientist. Specific examples of interdisciplinary researcher collaborations included conservation biologists, ecologists, population geneticists and behavioral scientists.

A key component of the Open Science movement relates to attitudes concerning the publication of work in progress. 67% of Scratchpad maintainers said they use their Scratchpad to publish work in progress, either privately to students, referees or colleagues, or publicly and completely openly on their site. When asked about the benefits of such activities, examples given include the advantages of getting feedback from other people, increasing the visibility of their work, attracting new people to a project and publishing small amounts of information which in themselves are not large enough for traditional paper publication. Several users also specifically referenced Open Science and Open Access benefits as part of their response, referring to the Scratchpad's utility in publishing data rich resources that are not possible in traditional paper publications. Users also cited the fact that particular categories of data (e.g., bibliographies and interactive identification keys) are by their nature works in progress, which are constantly refined and updated. Scratchpads support this functionality in a way that traditional publications cannot. Only 24 % of respondents stated that they would not publish works in progress. Reasons given include that this was not the purpose of their sites, and that some maintainers are weary of people stealing their data and are uncomfortable publishing untested hypotheses.

3.3 Technical and Social Barriers of Adoption

We asked interviewees about possible incentives and barriers in their work environment that facilitate, motivate or hinder use of their Scratchpads. An overwhelming number of respondents (85%) referred to their lack of time to maintain content (i.e. insufficient time in addition to their other duties to add, update and maintain site information). A significant number also cited their lack of time to acquire the technical skills necessary to develop their site. Further investigation suggested that a quarter of site maintainers lack general level computer literacy (25%) and cite this as a problem, rather than specific technical skills required to use the Scratchpad. Some maintainers were unsure of whether their time investment will be commensurate with rewards from using the site (20%), and upon further investigation this was mostly due to lack of trust in the technical development of Scratchpads. Just two maintainers had concerns that their data may be misused if published on their site and only one maintainer cited concern that they did not receive any credit from their employer for their time invested in site development. Follow up questions on institutional support for using new communication technologies as part of the respondents scholarly research revealed that a majority of maintainers (70%) received no institutional credit or incentives from their employer for their online work. However, the corollary of this is that the remaining 30% do receive credit for this work as part of their job evaluation. This picture broadly matches the pattern of institutional technical support received by these individuals in their use of these virtual research tools. 68% said they receive no technical support, 24% said they did, and a further 8% said they did not know. Arguably, this gap is filled by help and feedback facilities built into the Scratchpads. These include instructional video

screencasts and online forums that are integrated within individual users Scratchpads. 83% of respondents said they used these facilities and 46% said they were satisfied with them.

For many users a key impediment to their continued use of the Scratchpads was the fact that this tool is still in the development phase. 90% of maintainers think that Scratchpads need further improvement and only 7% were happy with current site functionality (2% did not know). The top three requests for enhancing the Scratchpads were i) improvements in the Scratchpads usability, specifically the sites need to be more intuitive; ii) the need for better tools to manage security levels for different members and different pages of the site; and iii) better management of taxonomic hierarchies. A regular comment was that the sites do not deal taxonomic classifications properly.

4 Conclusion

A primary motivation for traditional article publication is to demonstrate the authors' contribution to science. This attracts peer recognition that influences the authors' reputation, employment and research opportunities. Broadly speaking the survey reveals the same motivating forces are operating with the Scratchpads and in the on-line spaces created by other scholarly virtual research tools. The primary motivation for registering for a Scratchpad was the desire to collaborate over the Internet. In reality, many maintainers struggled to engage and sustain substantial research collaborations online. Consequently communication of information became the primary use for a majority of sites. These sites are in effect being used as digital surrogates for traditional research publications, and less so as tools to share or collaborate with data. Indeed, some maintainers actively cite concerns about sharing data as a possible barrier to continued use of the site. Experience by the developers has shown that many users try to add conditions of use to data on their sites that are restrictive or create obstacles to such use. This is despite the fact that Creative Commons licenses are enforced on site contributions as a condition of use. At the very least this suggest that many contributors do not understand Creative Commons licenses. This quasi-release of data by many Scratchpad users is not only contrary to goals of the Scratchpad project but also acts as a disincentive to others to explore the sites potential, or at least make use of the data.

These data suggest that motivating sustained and greater use of the Scratchpads and similar virtual research tools might be more effective if the traditional principles of publication, rather than just tools for data sharing, are built into the system. Article citation and journal impact factors are the most common metric of peer recognition and play an important role in evaluating the quality and impact of scientific research. The transition to new online forms of scholarly communication creates the potential for a rich and diverse set of new performance indicators that consider wider aspects of the process and are better able to predict and manage the outcomes of research. If a comparable metric could be brought to bear on content published in a Scratchpad, it follows that value of the online content could be similarly tracked to motivate authors.

Delivering methods to track online contributions is not technically easy, because they involve tracking diverse categories of contributions across multiple systems managed by many institutions and organizations for a potentially huge number of authors, each of which would need to be uniquely identified across these systems. Despite this challenge, some metrics such as the Scholar Factor (SF) proposed by Bourne & Fink [13] do provide a possible solution to this problem and should be investigated alongside methods for standardizing the citation and archival of online content. This approach would enable contributors to receive appropriate credit for their contributions and for peers to more effectively measure these contributions as part of a researcher's scholarly activities.

The survey revealed a number of technical improvements to the Scratchpads that were previously unknown to the Scratchpad development team. However, the primary barrier to use cited by the survey respondents was not technical, but relates to their lack of time to expand and develop the site. This was closely followed by a lack of technical skill in reaching the full potential of the Scratchpad application. A likely interpretation of this response is that the respondents struggled with assessing the value of engaging with these activities and the necessary changes to their working practices, especially with respect to novel ways of organizing their data outside traditional paper based publications. As was revealed by the survey, the use of virtual research tools is not part of most respondent's job evaluation and is not comparable with the value of traditional measures of peer recognition, such as obtaining papers in high impact journals. Top down incentives by managers will be required to change this perception, and this will require those managers to be convinced of the value in these online activities.

These data reveal that although some technical and social barriers do hinder uptake and use of the sites, even to the point that some users abandon their sites entirely, a majority strongly appreciate the possibilities that Scratchpads offer. These survey responses are backed up by site access logs which reveal that a considerable number of researchers have used their Scratchpads as part of their daily research routine, especially as a mechanism to deliver informal, rapid, open access publication of content. A smaller subset of users also reveal themselves as active collaborators, both off-line and online, operating in national and international research networks, and in some cases interdisciplinary networks. This strongly counters the traditional image of research taxonomists and reclusive loners, reluctant to engage in collaborative research activities.

The survey was a useful exercise for the Scratchpad developers in understanding the motivations, barriers and impact of the tool on their user community. However the ongoing agile development of the software coupled with constant refinements to the sites' functionality suggest this should not be a one-time exercise. Developments to the technology and continuous user feedback are necessary to cater for the diverse demands users place on the system. Study mechanisms are needed that are built into the Scratchpads. These need to be more agile, intensive and fun for users, rather than occasional labour intensive surveys such as the one reported here. The NHML Scratchpad team is leading a consortium that has recently obtained a further 3 years of EU funding, and as part of the project will develop online methodologies that embed these sociological study methodologies into the technology infrastructure. Our goal is

to obviate the need for similar survey activities and develop a mechanism that more rapidly informs the development team of user impacts and functionality needs.

Acknowledgements

The Scratchpad project has been financially supported by the [European Distributed Institute of Taxonomy](#) (EDIT), [The Natural History Museum](#) and [GBIF](#). We would like to thank Joanne Gerrits, Christine Hine and Simon Rycroft for their help with organizing the interviews and the interview script.

References

1. Suber, P.: Open Access Overview: Focusing on Open Access to Peer-Reviewed Articles and their Preprints. Available at <http://www.earlham.edu/~peters/fos/overview.htm> (2004)
2. Butler, D.: US Seeks to Make Science Free for All. *Nature* 464, 822-823, doi:10.1038/464822a. Published online 7 April 2010.
3. Neylon, C., Wu, S.: Open Science: Tools, Approaches, and Implications. *Pacific Symposium on Biocomputing*, 540-544 (2009)
4. Hull, D., Pettifer, S.R., Kell, D.B.: Defrosting the Digital Library: Bibliographic Tools for the Next Generation Web. *PLoS Computational Biology* 4(10), e1000204. doi:10.1371/journal.pcbi.1000204 (2008)
5. Murray-Rust, P.: Chemistry for Everyone. *Nature* 451, 648-651 (2008)
6. Lessig, L.: *Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity*. New York, NY, Penguin Press (2004)
7. Benkler, Y., Nissenbaum, H.: Commons-based Peer Production and Virtue. *The Journal of Political Philosophy* 14(4), 394-419 (2005)
8. Elliot, M., Scacchi, W.: Mobilization of Software Developers: The Free Software Movement. *Technology and People* 21(1), 4-33. Available at <http://www.emeraldinsight.com/Insight/ViewContentServlet?Filename=/published/emeraldfulltextarticle/pdf/1610210101.pdf> (2008)
9. Waldrop, M.: Science 2.0: Great New Tool, or Great Risk? *Scientific American*, available at <http://www.sciam.com/article.cfm?id=science-2-point-0-great-new-tool-or-great-risk> (retrieved on 24 November 2008)
10. De Roure, D.: Web 2.0 and the Democratisation of E-research. SSGC2008, Hungary, July. Available at <http://www.semanticgrid.org/presentations/issgcWeb3.ppt> (2008)
11. Smith VS, Rycroft SD, Harman KT, Scott B, Roberts D: Scratchpads: A Data-Publishing Framework to Build, Share and Manage Information on the Diversity of Life *BMC Bioinformatics* 10 (Suppl. 14): S6 doi:10.1186/1471-2105-10-S14-S6 (2009)
12. Knapp S.: Taxonomy as a Team Sport. In: Wheeler, Q.D. (ed.) *The New Taxonomy*. Pp. 33-53. CRC Press, New York (2008)
13. Bourne, P.E., Fink, J.L.: I Am Not a Scientist, I Am a Number. *PLoS Computational Biology* 4(12), e1000247. doi:10.1371/journal.pcbi.1000247 (2008)