

# Social Lab: An "Open Source Facebook"

Ulf-Dietrich Reips & Pablo Garaizar

The overlap between our every day activities and our behaviours on the Internet is ever increasing. With the advent of social media the social and behavioural sciences are faced with new opportunities and challenges for research into social behaviour. The vast majority of social media are owned by private companies. Despite public application programming interfaces (APIs) being offered by some of these social media, research in proprietary networks is severely limited. Considering the limitations to social media research, we have developed Social Lab, an open source clone of Facebook with most of its features (messaging, sharing, befriending, wall posts, pictures, searching, profiles, privacy settings, etc.). In addition, Social Lab enables researchers to create "social bots" – automated programmable profiles controlled through simple scripts – to facilitate the study of social phenomena. In the present chapter we introduce Social Lab using an example around privacy management in

social media, show how to configure social bots in Social Lab, and explain how it can be used in research. The source code of Social Lab is freely available to the scientific community, so any research group can have its own Social Lab to conduct their Internet-based research.

Keywords: Social media; Internet-based research; Internet science; Facebook; open source; social bots; Social Lab; social networks; big data; iscience

## INTRODUCTION

The Internet, "Web 2.0", and social media – the evolution of networked technologies and their social use – has yielded a new generation of researchers, who have developed a new set of methods and tools for research on the Internet and in its services (see e.g. Reips & Birnbaum, 2011; Reips & Buffardi, 2012).

Tools available for such research include: *FactorWiz* and *SurveyWiz* (Birnbaum, 2000); *iScience Maps* (Reips & Garaizar, 2011); innovative social location-aware services for mobile phones like *MUGGES* (Klein & Reips, this volume), *Scientific LogAnalyzer* (Reips & Stieger, 2004); the *Web experiment list* (Reips & Lengler, 2005); the *Web Experimental Psychology Lab* (Reips, 2001); *WEXTOR*, a Web experiment generator (Reips & Neuhaus, 2002); *ReCal OIR* (Freelon, 2013); *VAS Generator* (Reips & Funke, 2008); *Dynamic Interviewing Program* and *User Action Tracer* (Stieger & Reips, 2008, 2010); among many others. The number of studies conducted via the Internet with such tools has grown almost exponentially since 1995 (Reips & Krantz, 2010).

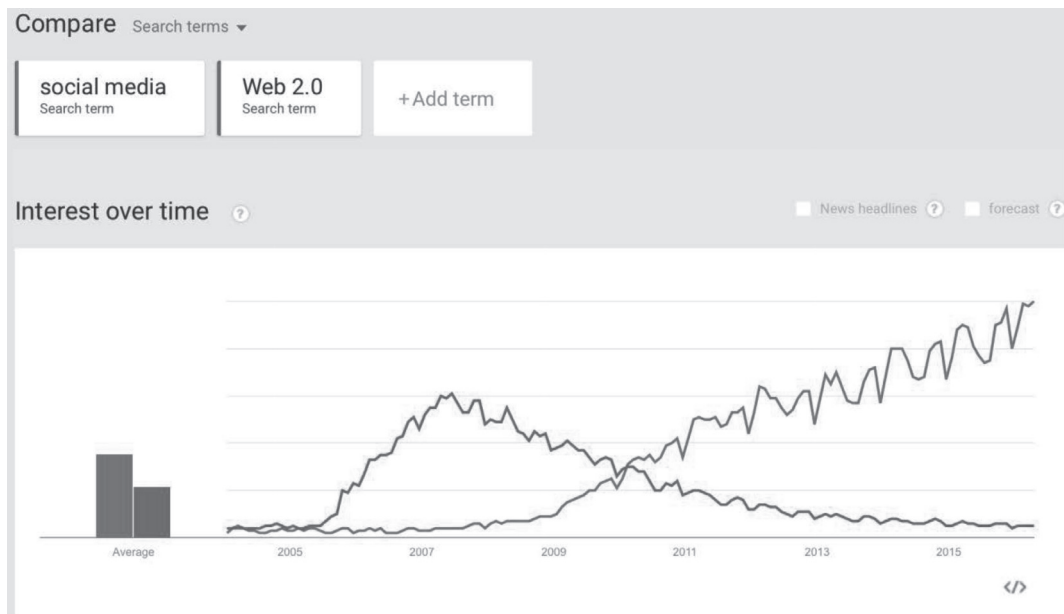
In the following sections, we will describe a new tool in this tradition, *Social Lab*, and how it can be used in research. We will also illustrate its application with an example, in which Social Lab is used to learn and practice privacy management in Social Networking Sites (SNSs). We will then

provide instructions on how to extend its functionality for particular research purposes by configuring so-called “social bots”, automated agents within Social Lab that to the user may often appear indistinguishable from human users of the social network.

## SOCIAL MEDIA RESEARCH OPTIONS AND BIG DATA

“Social media is a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user generated content” (Kaplan & Haenlein, 2010, p. 61). Figure 27.1 empirically shows with a Google Trends analysis how the term “social media” became more popular than “Web 2.0” and continues to be on the rise (also see McCay-Peet and Quan-Haase, this volume).

Unknown to many people, who may just know about the largest social media sites



**Figure 27.1** Analysis of searches for the terms “Web 2.0” and “social media” over time in Google Trends

like Facebook, Twitter, Instagram, Snapchat, Weibo, LinkedIn, Orkut, Tuenti, Google+, Wikipedia, Tumblr, or YouTube, there are thousands more of social media that often aim at particular interests or groups. For example, Wikipedia lists almost four hundred “major active” sites (Wikipedia, 2016), including many “second level” SNSs for special, albeit large, interest groups, for example, academics (ResearchGate, Academia.edu, etc.). Social media can be categorized in various ways, for example, by main user function (publish, share, discuss, livestream, microblog, live-cast, play in virtual world or a social game) or industry purpose (social advertising, marketing, analytics, streaming, social data mining, social intelligence, social scoring, internal or external business software, reviewing, social shopping, social referral, content curation, social TV, social brand engagement, etc.).

Research using social media can be conducted on many topics of social interaction that don't require physical co-presence or a synchronous face-to-face situation. Social phenomena have been shown to appear online as well as offline, see, for example, research on ostracism (Vorderer & Schneider, in press) or migration (Reips & Buffardi, 2012; Oiarzabal & Reips, 2012). Communication in social media and the effects of social media use (e.g., emotional consequences of using Facebook [Lin & Utz, 2015], or even Facebook addiction [Dantlgraber et al., in press]) have become a topic of interest and research in itself. Both the research on social media use and on how social media can be used for research purposes have increased enormously during the last few years (also see McCay-Peet and Quan-Haase, this volume).

Because very large SNSs like Facebook have been frequently involved in research, data sets tend to be very large. Big Data has thus become an important topic in social media research and Internet science, with related issues. Big Data gathered from social media may in fact provide answers to one of the main issues in behavioural and social science research with very large data sets, the

lack of connection between the micro and macro levels of analysis (Snijders et al., 2012).

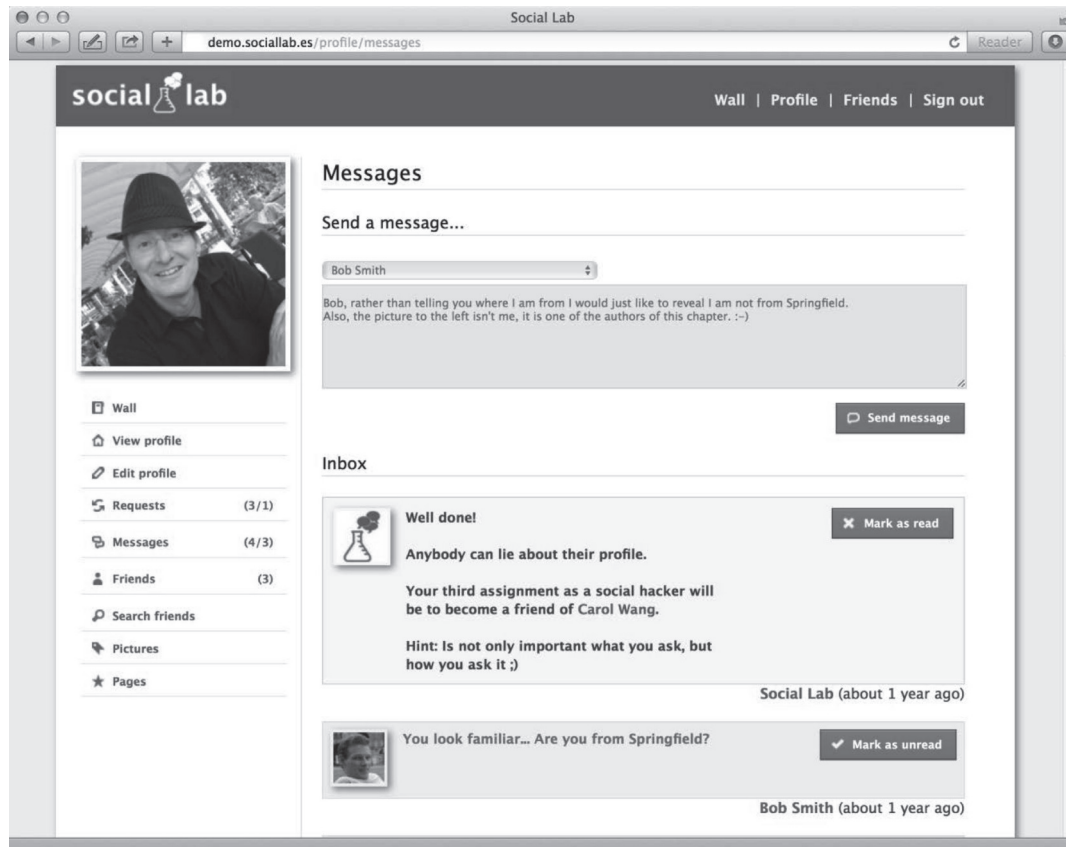
To select a social media site for research or other purposes, one can use a tool called Social Media Planner (<http://www.inpromo.de/wordpress/en/social-tools/social-media-planner/>). In a first step, one selects demographics and interests, and in a second step, the tool generates a list of social media sites that are geared toward the selected criteria (e.g., education) and are used by (mostly) people with the selected range of demographics.

On the downside, existing SNSs come with many disadvantages related to their often proprietary nature (ads; lack of true access to the data; end of service due to economic issues like buyout, bankruptcy, change of business; business rather than research driven structure of data; change of service, etc.) and sampling biases. Even though his methodology of using Facebook likes may be somewhat questionable with respect to details of the empirical results, Ruffini (2012) convincingly describes how the selection of a social media service may lead to a particular, possibly biased sample.

## SOCIAL LAB: WHAT IT IS

Social Lab can be described as an open source “Facebook” clone, a fully functional and free software to run social networks for research and other purposes (Garaizar & Reips, 2014). Social Lab is available from <http://www.socallab.es> and provides most features commonly found in other social media, for example, messaging, sharing, befriending, wall posts, pictures, searching, profiles, privacy settings, etc. (see Figure 27.2).

Having full availability of navigation and communication data in Social Lab allows researchers to investigate behaviour in social media on an individual and group level. Automated artificial users (“social bots”, for a similar concept of *socialbots* not native to a SNS see, e.g., Boshmaf et al., 2011) are



**Figure 27.2 Main display of Social Lab**

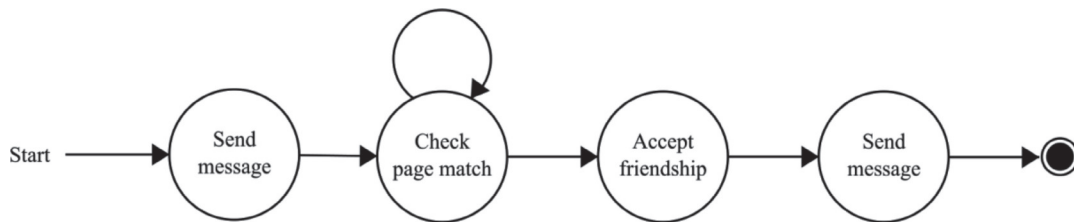
available to the researcher to simulate and stimulate social networking situations. These bots respond dynamically to situations as they unfold and they have a memory for their previous interactions. The social bots can easily be configured with simple scripts and can be used to experimentally manipulate social networking situations in Social Lab.

### **AN EXAMPLE: AN EDUCATIONAL PRIVACY MANAGEMENT SITE USING SOCIAL LAB**

To exemplify how Social Lab can be used, we set up a Social Lab site that helps users develop a sense of the challenges around privacy management in social media. Privacy

on the Internet has been and continues to be a major issue (Joinson et al., 2010; Reips, 2011). The Social Lab privacy education site is available in English at <http://en.sociallab.es/>, in Spanish at <http://es.sociallab.es/>, in German at <http://de.sociallab.es/>, and in Basque at <http://eu.sociallab.es/>. Its purpose is to demonstrate some of the techniques used by social hackers in order to enable users to prevent this kind of attacks in real social networks.

Anyone can sign up and get started with this site. Once signed in, the first challenge in protecting users' privacy is waiting in the inbox. Every challenge will be controlled by an automated, script-based profile that appears like a fellow user on the social network. Profiles must be convinced to become friends using one's social engineering skills.



**Figure 27.3 Simple social bot for Social Lab**

Their responses are automatic but may not be immediate. Sometimes they need some time to respond, just like real users.

Figure 27.3 shows how the site provides feedback to users. As we mentioned before, the ten challenges in the privacy management example site and their associated profiles are based on social bots, an essential feature of Social Lab. In the next section we will explain how to create and program bots in Social Lab.

## PROGRAMMING A BOT IN SOCIAL LAB

From a technical perspective, Social Lab’s bots are regular user accounts provided with unattended behaviours. Therefore, researchers can make a social bot from a user account previously managed by a person. The opposite is also feasible: a person can take manual control of an account previously managed by a social bot. Moreover, sharing the management of the user account between a person and a social bot at the same time is also possible. This flexibility in managing social bots’ accounts enables a wide variety of research scenarios.

The first step to program a bot on Social Lab is to create a user account and define its profile (i.e., first name, last name, email, gender, birthday, picture, location, academic information, personal information, privacy options) using the registration page (<http://yoursociallab.com/signup>). We use

“yoursociallab.com” as a fictitious server name where Social Lab is deployed by a research group. Most of these features are in use in our online privacy game example. Therefore, the reader can replace “yoursociallab.com” with “en.sociallab.es” to see them working. The rest of the process of defining the automatic behaviour of the social bot is done through the Social Lab backend interface (i.e., <http://yoursociallab.com/backend.php> – for security reasons this URL won’t work in our example at <http://en.sociallab.es> to prevent hackers from accessing our privacy game backend).

In Social Lab the behaviour of a social bot is defined by a sequence of steps. There are two types of steps: actions and checks. *Actions* (e.g., send message, accept friendship request, etc.) are executed in a predefined order and always enable the processing of the next step. *Checks* (e.g., check if someone is a friend of two friends of mine, check location match, etc.), by contrast, are evaluated whenever they are requested by the social bot, but they do not allow to process the next step unless their condition is met.

Social bots keep in the database the status of each interaction with each user. Thus, they may be running step 1 with the user A while performing step 3 with the user B. When a social bot has executed all the steps defined in its behaviour interacting with a given user, this interaction ends. However, interactions with other users will run their course in the state in which they were.

Let’s see how to define the behaviour of a simple bot. As shown in Figure 27.3, this bot

will send a message asking the user whether she likes Mozart or not in the first step. In the second step, the bot will wait until the user it is interacting with adds the Mozart page to her preferences. The bot will wait in this state indefinitely and will not execute the rest of its program until the condition is met. When that is the case, the bot will accept the user's friendship request (third step) and send a message recommending other pages of similar musicians (fourth step).

Defining this automatic behaviour in Social Lab is straightforward. First, a regular user account is created, it must then be added to the "Bots" table using the backend application (<http://yoursociallab.com/backend.php/bot/new>). Then, all the steps defined in Figure 27.3

have to be added to the "Steps" table (<http://yoursociallab.com/backend.php/step/new>), setting the following parameters, see Figure 27.4: 1) the bot that we are programming (a drop-down list of all bots defined in the "Bots" table is shown here to minimize input errors); 2) the command to be executed in this step ("Send message" for the case of the first step of this social bot); 3) the step order (0 for the case of the first step); and 4) a reference to the message intended to be sent. All text messages sent by social bots have to be stored in the "Automsg" table. Therefore, it is necessary to add a new entry to this table (<http://yoursociallab.com/backend.php/automsg/new>), indicating the text to send. The second step of the social bot has to be added in a similar

**EDIT STEP**

|            |                       |
|------------|-----------------------|
| Bot        | Level 2: Ask location |
| Command    | Send message          |
| Step order | 0                     |
| Automsg    | You look familiar...  |
|            |                       |
| Bot        | Level 2: Ask location |
| Command    | Check location match  |
| Step order | 1                     |
| Automsg    |                       |
|            |                       |
| Bot        | Level 2: Ask location |
| Command    | Accept friendship     |
| Step order | 2                     |
| Automsg    |                       |
|            |                       |
| Bot        | Level 2: Ask location |
| Command    | System message        |
| Step order | 3                     |
| Automsg    | Well done! <br/><br/> |

**Figure 27.4** Setting up triggers and routines for a social bot in Social Lab's backend application

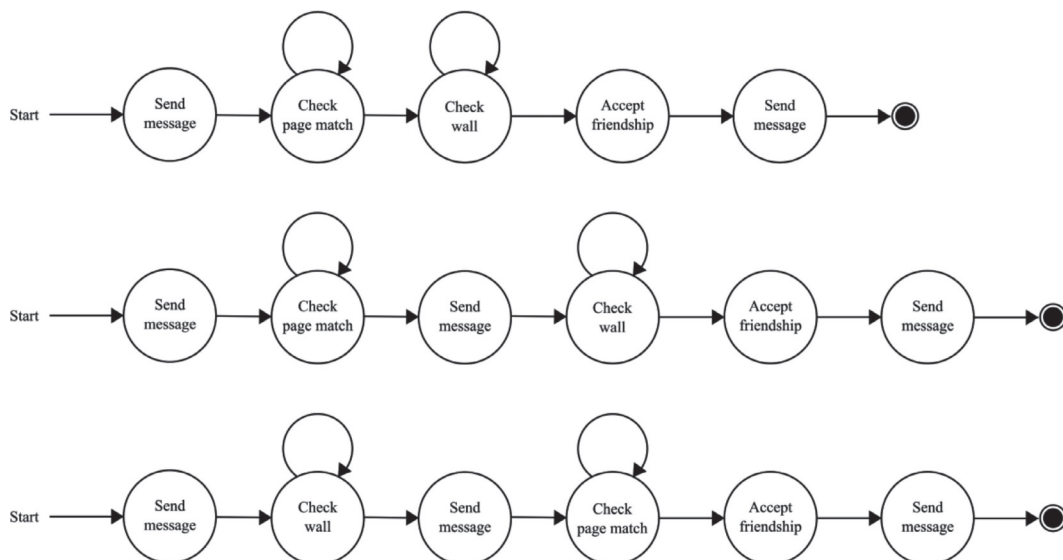


way, adding a new entry to the “Steps” table (<http://yoursociallab.com/backend.php/step/new>). In this case, the command should be “Check page match”, the step order should be 1, the automsg field should be left blank because this check does not require a message. The third step is defined in the same vein, adding a new step for the same bot and defining the command (accept friendship) and the step order (2). Finally, the last step of creating this bot involves another “Send message” command. Therefore, the message to send should be added previously to the “Automsg” table (<http://yoursociallab.com/backend.php/automsg/new>). Once this procedure is finished, the bot is ready to interact with users of the social network.

As we can see, defining a bot in Social Lab does not require previous technical knowledge, just adding values to the database. Thanks to the backend application provided with Social Lab, these database changes can be done with a form-based interface. With this social bots definition system a researcher could easily compare two similar situations by creating two social bots with slightly

different behaviours and analyzing, which interactions occur with each of them. For example, in a study related to participation in online promotions researchers might want to know whether participants asked to perform two actions (e.g., mark a page as favourite and post a message on their wall) to take part in the online promotion are more reluctant to participate than those who first are asked to perform one action first and then the other one. Researchers conducting this experiment could define three bots as shown in Figure 27.5 and compare the number of users in the social network that reach the last step in each case. Comparisons between the two last bots could also give clues about the reluctance associated with each action.

User accounts of each of these bots can also be simultaneously managed by humans. Therefore, the community manager responsible for encouraging participation in this online promotion could use the social bots user accounts manually in some moments of the marketing campaign. This increases the feeling of interacting with a person and not a program, which could also be the subject



**Figure 27.5** Three slightly different social bots to gather participants for an online promotion

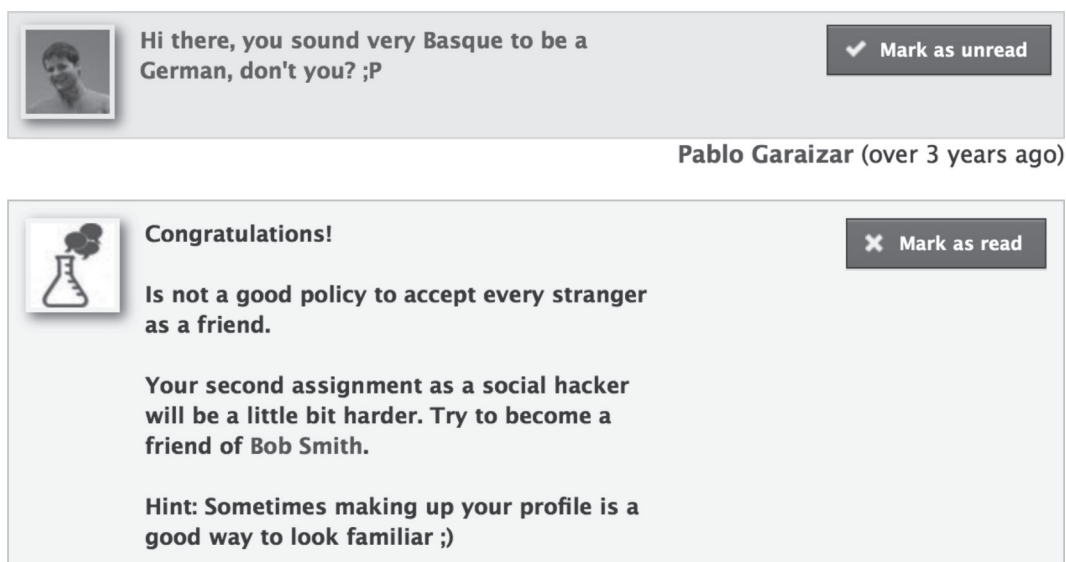
of another related study (i.e., compare the engagement in online promotions boosted by profiles clearly run by people and by profiles clearly managed by bots). As mentioned before, this shared management of user accounts also enables multi-phase studies in which the behaviour (e.g., bot A, bot B and bot C in Figure 27.5) and type of interaction (manual vs. automated) changes in each phase but all interactions are made from the same user account.

Moreover, it is also possible to take advantage of the fact that all behaviours of all social bots are defined as rows of a small set of tables in the Social Lab database (i.e., bot, step, automsg). Thanks to this decoupled design, third-party software could dynamically reprogram the behaviour of the bots modifying the content of those tables without modifying a single line of the particular Social Lab instance's source code. This enables the integration of sophisticated behaviours (e.g., chat bots managed by complex Artificial Intelligence algorithms) without having to integrate their code in Social Lab, simply making changes in the database.

## RESEARCH OPTIONS WITH SOCIAL LAB

Here we outline ideas about how Social Lab could be used in research. In the most simple study design, researchers can simply observe how users act and communicate within the social network. Like in Figure 27.6, messages can be deployed by bots or researchers to provoke user action. As fully controllable open source software, Social Lab lets researchers access more information than any proprietary network.

Also, researchers may use Social Lab to investigate group-level social network structure and dynamics. Social network analysis could then be used in the analysis of the social network – for example, research into changes of the pattern of connections between persons in a collective (Quan-Haase & McCay-Peet, 2016). Of course, a network set up with Social Lab can provide an environment for questionnaire studies and interviews as well. Such interviews could even be conducted automatically using a pre-programmed bot that randomly contacts users within the



**Figure 27.6** Messages can be deployed by researchers or social bots or to provoke user actions



social network and invites them to participate (Stieger & Reips, 2008).

One could further imagine *experimental* studies that are based on Social Lab. So far, to our knowledge there has been only one truly experimental study with Internet-based social networks (Centola, 2010), i.e. randomly assigning participants to different social networks. Centola (2010) was interested in the influence of the structure of social networks on the spread of behaviour and randomly assigned participants who had signed up to a health forum to either a clustered lattice network condition and a random network condition. Participants were further randomly assigned to a node in their network, i.e. the other participants they would see within the network were determined by this assignment. Further, each participant could receive at most one email from each of the participants they could see and they could not contact each other within the social network. Improving on this communication-poor, rather static and purely structural manipulation one could use the human-like bots available in Social Lab to for example influence risk perception in two initially structurally identical networks and observe risk communication, coping behavior, and attitude change. In pre-study and follow-up surveys one could collect attitudinal information and self-reported behaviors and emotions from participants and then see how these change depending on the experimental manipulations and resulting dynamics within the networks.

*Setting up Social Lab as an Experimental Environment.* Experiments can be conducted *within* a social network or *between* social networks. In the latter case, at least two structurally identical social networks need to be created using Social Lab after download and setup from <http://www.sociallab.es>. Participants should be recruited from different Web sites and in different modes (e.g., actively: “push” or passively: “pull”). Via the multiple site entry technique (Reips, 2000, 2007) the researcher will be able to determine for each participant the site he or she

was recruited from and will thus later be able to test for systematic differences between recruitment sites and modes. For example, one could aim for 2,000 participants (ca. 500 in each network), then have them fill in informed consent forms and randomly assign them to one of the four social networks. The social networks then need to be closed; no new applications will be accepted during the time of the studies.

Social Lab will often need to be adjusted to meet the requirements of a particular study. For example, there may be a need to limit configuration options (e.g., delete the informed consent option within Social Lab, if informed consent was already collected before the logon). One could also amend options to display “news items” that can be used for experimental manipulations. As a way to manipulate exposure (e.g., via frequency and intensity of messages or via number of agents) and role modeling of coping strategies, several social bots can be created in all networks and their actions and timing will be programmed to serve the tasks required by the experimental manipulations.

## OUTLOOK

We have presented the free open source social network software Social Lab and outlined how it can be used in research. Social Lab provides most features that are common to social media and even more: one of the most useful features in research are its automated programmable bots. We described above how such social bots can be used in research and how to create and configure them within Social Lab.

Much research including social network analysis can be applied to social networks that are developed within Social Lab or manifest themselves in communications on this platform. With Social Lab, researchers can collect information on who is connected to and communicates with whom, in time. Such

data can then be used to investigate social network properties, such as density, clustering, and connectedness, or node properties, such as betweenness or centrality (see e.g., Brandes & Erlebach, 2005; D'Andrea et al., 2010).

We hope Social Lab will inspire many to set up free social networks for research and beyond. An open source community has begun to form and help in its development. May the new tool help the scientific community to advance in the exciting study of behaviour in social media.

## REFERENCES

- Birnbaum, M. H. (2000). SurveyWiz and FactorWiz: JavaScript Web pages that make HTML forms for research on the Internet. *Behavior Research Methods, Instruments, & Computers*, 32(2), 339–346.
- Boshmaf, Y., Muslukhov, I., Beznosov, K., & Ripeanu, M. (2011). The socialbot network: when bots socialize for fame and money. In *27th Annual Conference on Computer Security Applications* (p. 93–102). <http://doi.org/10.1145/2076732.2076746>
- Brandes, U., & Erlebach, T. (Eds.) (2005). Network analysis: Methodological foundations. *Lecture Notes in Computer Science Tutorial*, vol. 3418. Berlin: Springer.
- Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, 329, 1194–1197.
- D'Andrea, A., Ferri, F., & Grifoni, P. (2010). An overview of methods for virtual social network analysis. In A. Abraham, A.-E. Hassanien, & V. Snasel (Eds.), *Computational social network analysis: trends, tools and research advances* (pp. 3–26). Berlin, Germany: Springer. doi:10.1007/978-1-84882-229-0
- Dantlgraber, M., Wetzel, E., Schützenberger, P., Stieger, S., & Reips, U.-D. (in press). Simple construct evaluation with latent class analysis: An investigation of Facebook addiction and the development of a short form of the Facebook Addiction Test (F-AT). *Behavior Research Methods*. doi:10.3758/s13428-016-0716-2
- Freelon, D. (2013). ReCal OIR: Ordinal, interval, and ratio intercoder reliability as a web service. *International Journal of Internet Science*, 8(1), 10–16.
- Garaizar, P., & Reips, U.-D. (2014). Build your own social network laboratory with Social Lab: A tool for research in social media. *Behavior Research Methods*, 46(2), 430–438.
- Joinson, A. N., Reips, U.-D., Buchanan, T., & Paine Schofield, C. (2010). Privacy, trust, and self-disclosure online. *Human-Computer Interaction*, 25, 1–24. doi: 10.1080/07370020903586662
- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53(1), 59–68. doi:10.1016/j.bushor.2009.09.003
- Lin, R., & Utz, S. (2015). The emotional responses of browsing Facebook: Happiness, envy, and the role of tie strength. *Computers in Human Behavior*, 52, 29–38. doi:10.1016/j.chb.2015.04.064
- Oiarzabal, P. J., & Reips, U.-D. (2012). Migration and diaspora in the age of information and communication technologies. *Journal of Ethnic and Migration Studies*, 38(9), 1333–1338. doi:10.1080/1369183X.2012.698202
- Quan-Haase, A., & McCay-Peet, L. (2016). Social network analysis. In K. J. Bruhn & R. T. Craig (Eds.), *International encyclopedia of communication theory and philosophy*. Cambridge, MA: Wiley.
- Reips, U.-D. (2011). Privacy and the disclosure of information on the Internet: Issues and measurement. In A. Blachnio, A. Przepiórka & T. Rowinski (eds.), *Internet in Psychological Research* (pp. 71–104). Warsaw: UKSW Publishing House.
- Reips, U.-D. (2007). The methodology of Internet-based experiments. In A. Joinson, K. McKenna, T. Postmes, & U.-D. Reips (Eds.), *The Oxford Handbook of Internet Psychology* (pp. 373–390). Oxford: Oxford University Press.
- Reips, U.-D. (2001). The Web Experimental Psychology Lab: Five years of data collection on the Internet. *Behavior Research Methods, Instruments, and Computers*, 33, 201–211.
- Reips, U.-D. (2000). The Web Experiment Method: Advantages, disadvantages, and

- solutions. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 89–118). San Diego, CA: Academic Press.
- Reips, U.-D., & Birnbaum, M. H. (2011). Behavioral research and data collection via the Internet. In K.-P. L. Vu & R. W. Proctor (Eds.), *The handbook of human factors in Web design* (2nd ed., pp. 563–585). Mahwah, New Jersey: Erlbaum.
- Reips, U.-D., & Buffardi, L. (2012). Studying migrants with the help of the Internet: Methods from psychology. *Journal of Ethnic and Migration Studies*, *38*(9), 1405–1424. doi:10.1080/1369183X.2012.698208
- Reips, U.-D., & Funke, F. (2008). Interval level measurement with visual analogue scales in Internet-based research: VAS Generator. *Behavior Research Methods*, *40*, 699–704.
- Reips, U.-D., & Garaizar, P. (2011). Mining Twitter: Microblogging as a source for psychological wisdom of the crowds. *Behavior Research Methods*, *43*(3), 635–642.
- Reips, U.-D., & Krantz, J. H. (2010). Conducting true experiments on the Web. In S. Gosling & J. Johnson (eds.), *Advanced methods for conducting online behavioral research* (pp. 193–216). Washington, DC: American Psychological Association.
- Reips, U.-D., & Lengler, R. (2005). The Web Experiment List: A Web service for the recruitment of participants and archiving of Internet-based experiments. *Behavior Research Methods*, *37*, 287–292.
- Reips, U.-D., & Neuhaus, C. (2002). WEXTOR: A Web-based tool for generating and visualizing experimental designs and procedures. *Behavior Research Methods, Instruments, and Computers*, *34*, 234–240.
- Reips, U.-D., & Stieger, S. (2004). Scientific LogAnalyzer: A Web-based tool for analyses of server log files in psychological research. *Behavior Research Methods*, *36*(2), 304–311.
- Ruffini, P. (2012). Infographic: Mapping the politics of the Social Web. <http://enga.ge/design/mapping-politics-social-web/>
- Snijders, C., Matzat, U., & Reips, U.-D. (2012). 'Big Data': Big gaps of knowledge in the field of Internet science. *International Journal of Internet Science*, *7*(1), 1–5.
- Stieger, S., & Reips, U.-D. (2008). Dynamic Interviewing Program (DIP): Automatic online interviews via the instant messenger ICQ. *CyberPsychology and Behavior*, *11*, 201–207.
- Stieger, S., & Reips, U.-D. (2010). What are participants doing while filling in an online questionnaire: A paradata collection tool and an empirical study. *Computers in Human Behavior*, *26*, 1488–1495.
- Vorderer, P., & Schneider, F. M. (in press). Social media and ostracism. In K. D. Williams & S. A. Nida (Eds.), *Social exclusion*. New York, NY: Psychology Press.
- Wikipedia (2016, April 16). List of social networking websites. Retrieved April 21, 2016, from [https://en.wikipedia.org/w/index.php?title=List\\_of\\_social\\_networking\\_websites&oldid=715618422](https://en.wikipedia.org/w/index.php?title=List_of_social_networking_websites&oldid=715618422)