Bridging Social Media Technologies and Scientific Research: A Twitter-Enabled Platform for VPH Modeling

Vangelis Sakkalis, Stelios Sfakianakis, and Kostas Marias

Computational Medicine Laboratory, Institute of Computer Science, FORTH, Heraklion, Greece
{sakkalis,ssfak,kmarias}@ics.forth.gr
http://www.ics.forth.gr/cml/

Abstract. Social media and the Web2.0 technologies are ubiquitous and due to the advances in mobile communication protocols, operating systems, and internet standards they are now supported even in cell phones and tablets. We are not yet at the point where a cell phone can be used as a medical device but such small and omnipresent instruments can be used in a way that promotes research in the clinical and biomedical domain. In this paper we describe a collaborative platform for designing composite simulations for the Virtual Physiological Human (VPH) community needs. We investigate the use of pervasive mobile technologies so that scientists and researchers can easily design, share, and execute simulations. The proposed platform supports real time notification and sharing of the results, and share the results and related artifacts with their work group and colleagues.

Keywords: Social Networks, inSilico Oncology, Scientific Workflows.

1 Introduction

The focus of this paper falls in healthcare and more specifically in the research VPH community [1] [2] . Social web 2.0 and especially Medicine 2.0 refer to the generation of content by users, the power of networks, personalized health care, and the focus on collaboration across all stakeholders [3].

Starting back at the late 90's there have been a strong initiative to effectively apply Social Media tools to improve Heath care by actively engaging apart from patients themselves, researchers, practitioners and whole hospitals with the mission to promote health and fight disease. On one hand it is difficult enough to engage a busy physician and even more, specialized staff of hospitals and research centers. Fortunately, on the other hand, social media is the vehicle to decrease diffusion time for cancer research and innovations. Both patients [4] and researchers can benefit from such an empower network in very specific ways. Patients, in sites like PatientsLikeMe[1], may share intimate details of their

[1] http://www.patientslikeme.com

B. Godara and K.S. Nikita (Eds.): MobiHealth 2012, LNICST 61, pp. 380–387, 2013.

symptoms, diagnosis and treatment as well as discuss in respective groups about their condition [5]. This becomes even more crucial in diseases like cancer where treatment is merely personalized [6] and can be a controversial subject. In parallel, the research community may advocate for and against certain therapeutical schemes and computational "inSilico" models [7] [2] by rating and sharing their views on published models. Of course such online exchanges might not be ideally documented, but since they engage only experts in the field of computational oncology it is expected to disseminate accurate and noteworthy information that could pave the way towards new effective therapies.

In this paper we argue that the scientific community can gain a lot by the adoption of the social media tools and practices. As a show case we focus on the integration of a web application for the construction of scientific workflows with the micro-blogging service Twitter. In the following sections we present our application and the relevant scientific domain of interest. Then we describe the Twitter platform and we use its social networking infrastructure for the VPH modeling needs. In conclusion we present the challenges, the benefits of the approach, and future enhancements.

2 The Social Web

In spite of its success and popularity the early version of the World Wide Web lacked in many respects, ranging from user accessibility and user interface design, to the ability to repurpose and remix existing Web-based data in not pre-established ways. Most of these concerns have been addressed by advanced technologies for searching (Search engines), syndicating (RSS feeds), visualization and presentation mechanisms (CSS, SVG), etc. An additional limitation of this environment is that the people are not part of the equation. Users are expected to be the actors triggering the web interactions but content delivered should be personalized, relevant to the users context and needs, and users communication and collaboration should be promoted.

These and other requirements are the ones that the Social Web tries to tackle. Social Web does not represent a shift or radical change in technology per se but it represents rather a shift on the perception of the human – machine interaction by placing the users in the centre of the system and in control of these interactions. The requirements for implementing Social Web led to the emergence of a new breed of web applications and sites, collectively identified as "Web 2.0" by Tim O' Reilly [8], whose major design principle is to "harness network effects to get better the more people use them". The value of "Web 2.0" sites and applications therefore comes to a large extent by the number of users participating and actively communicating and sharing through them so the term "Social Web" is actually a synonym. The social nature of this Web is evident when the collaboration of people and their active contribution is considered.

Social Websites require a Social Network Infrastructure enabling of building social relations among individual researchers. Such a infrastructure provides services to represent a user by a personal profile, to develop social networks with

other actors, and to generate, collect and aggregate social activity in an activity stream. The main function is to create and update User Description profiles, to develop *follower networks*, to interact in the social environment, and to build and access activity streams. Users store personal data in user profiles and publish the data in their social network. In their follower networks, users can follow the activity of other users, activity in groups and communities, or on "tags". In the next paragraph we describe Twitter, one such social network infrastructure, that we integrate with.

2.1 Twitter

Twitter is a form of free micro-blogging which allows users to send and receive short public messages called "tweets". Tweets are limited to no more than 140 characters, and can include links to blogs, web pages, images, videos and all other material online. By following other people and sources, users are able to build up an instant, personalized Twitter feed that meets their full range of interests, both academic and personal.

More than 200 million are the daily Twitter users worldwide including thousands of academics and researchers at all levels of experience and different disciplines. The hard limit of 140 characters for the tweets provides the benefit of more instant updates and better "throughput" in terms of the number of messages a user is expected to read. Especially in the research community social media seems to surpass traditional publication media: Papers are increasingly being taken apart in blogs, on Twitter and on other social media within hours rather than years, and in public, rather than at small conferences or in private conversation [9] . Twitter forms social (sub)networks of people with common intentions and interests. Such networks were found to have a high degree correlation and reciprocity, indicating close mutual acquaintances among users [10].

A Glossary of Frequently Used Twitter Terms

Tweet (noun). A message posted via Twitter containing 140 characters or fewer.

Follow. To follow someone on Twitter is to subscribe to their Tweets or updates on the site. Following another user means that all their tweets will appear in your feed.

Follower. A follower is another Twitter user who has followed you.

Hashtag. The # symbol is used to mark keywords or topics in a Tweet.

Direct Message. A tweet that is private between the sender and recipient.

Mention. Mentioning another user in your Tweet by including the @ sign followed directly by their username is called a "mention"

Retweet (noun). A tweet by another user, forwarded to you by someone you follow. Often used to spread news or share valuable findings on Twitter.

3 A Web Based Platform for Scientific Workflow Design

A *workflow management system* is a computer system that manages and defines a series of tasks within an organization to produce a final outcome. In essence, a workow can be abstracted as a composite service, i.e. a service that is composed by other services that are orchestrated in order to perform some higher level functionality. In the recent years, Scientific workflow systems have emerged as the enabling technology for e-Science. Scientists collaborate on large scale scientific experiments and knowledge discovery applications using distributed systems of computing resources, data sets, and devices [11],[12].

The Thespis Workflow system is a web based graphical workflow designer focusing on the construction of VPH "hypermodels" by connecting "simpler" (atomic) models. The objective of this tool is to provide an intuitive environment for biomedical researchers and computational biologists where simulation models retrieved from different model repositories [13], can be combined to form higher level experiments implemented and managed as "workflows". Hence, Thespis is a useful platform for linking different models reflecting multiple biocomplexity levels in order to simulate complex processes (e.g. "microscopic" for genes and enzymes, to "macroscopic" scales at the tissue level) [14]. Figure 1 illustrates a typical workflow where different simulation models and are linked together. Through the exchange of data via the connections between their inputs and outputs perform a complex high level task. Such workflows can be the tools for the implementation of multi-scale biological modeling [7].

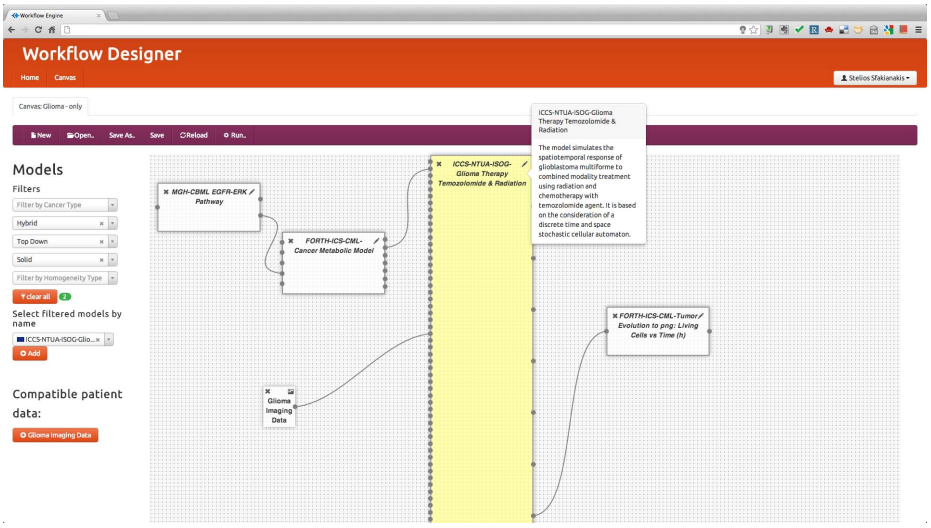

Fig. 1. A screenshot of the workflow designer

The Thespis workflow environment consists of two main components:

- The **workflow editor** (or designer) is a web application, accessible through the users' web browser. This is the graphical front-end for the editing of the workflows, the invocation of their execution, and the visualization of the results.
- The **workflow engine** is the server side, which is responsible for the management and the execution of the workflows, the communication with the model repositories, etc.

From the user point of view each user is required to register with the application and provide some information about himself, like the organization he works with, a description of his research and other interests, etc. The user authentication is supported either in a "traditional" way through some user name and password, or by linking their account with the account they maintain in some social networking site like Twitter, Facebook, or Google. In the latter case the authentication is taking place in the social networking site and the workflow application is granted authorization using the open standard oAuth [15]

In this application one can generally recognize three different types of users:

- The Model Creator (MC), who is the VPH researcher that builds new models and make them available, by publishing them in some model repository [13].
- The Workflow Creator (WC), who is the researcher that designs a new workflow by selecting existing models and tools, coupling them together based on their inputs and outputs, and possibly providing fixed parameter values for some of them.
- The Workflow User (WU), who is the user running ("enacts") existing workflows by providing some input values if needed, monitors their execution, and retrieves their results.

It is obvious that this classification of users does not represent mutually exclusive user groups. For example, it can be the case that a workflow user is also a workflow creator or that a workflow creator is also a model creator. In fact, the latter case corresponds to a future feature of the presented web application where a workflow can be reused and made available as a new model, albeit a "composite" one.

3.1 Social Based Facilities and Interactions

In the core of the user management module of the workflow application there are a number of Social Web inspired mechanisms that aim to enable the collaboration and communication among the users:

- The workflows can be annotated with "tags", i.e. free text keywords supplied by the users. These keywords can be used for the searching and classifying the workflows. By default the user submitted tags are public, which means that they are visible and can be used as filters in other users' searches, but if desired the user can make them private. Additionally private and public tags can be intermixed in the annotation of a single workflow.

- The users can have "favorite" workflows. This is useful to the workflow users so that they can better classify the most valuable workflows for their research. It is also useful to the workflow creator in order to get some feedback about the use of his/her workflows and possibly increase their "reputation" points.
- The platform keeps also statistics about the use of the workflows and the models in order to be able to answer questions like what is the most frequently used model or the workflow that has been enacted the most times.

However, to further leverage the potential of the social networks and especially Twitter we have defined a number of scenarios that reuse existing social networks and the online activity of the users. The objective of this work is to bring researchers together and to provide facilities for improving and advancing their work. To this end we have identified the following use cases where integration with a social platform like Twitter can help:

- Workflow creators and users want to know about the usage and the rating of their workflows and models, respectively.
- Workflow users want to notified when a new version of their favorite workflow is available.
- Workflow creators want to be notified when a new version of a model participating in their workflows is released.
- All users may be interested for new models become available.

For these and other scenarios the Twitter infrastructure is used to provide status updates for the interested users. First of all the workflow application maintains a twitter account (`@thespisapp`) as a "broadcast" account for news, notifications, etc. The users are encouraged to follow this user in order to get updates for new features of the application, general news, etc., but also for being able to receive personalized "direct messages" from the application. Furthermore, Thespis is using tweets for information that is potentially interesting to all its followers and direct messages if the information is considered more personal or of limited scope. Additionally it uses "hash tags" to annotate the messages with keywords that can be used for filtering and defining a personalized direct messaging functionality of Twitter. Examples of how this functionality is implemented through Twitter follows:

- When a new model is registered `@thespisapp` *tweets* about it *mentioning* the twitter user who created the model.
- When a new workflow is created `@thespisapp` *tweets* about it *mentioning* the twitter user who created the workflow and additionally sends *direct messages* to the creators of the models participating in this workflow.
- When a workflow is altered, the users who have annotated it as "favorite" receive *direct messages*.
- When a workflow becomes "favorite", its creator is notified with a *direct message mentioning* the twitter account of the user who liked it.

In general the methodology is to keep the status updates of the broadcast account of the application "noise-free" and prefer the use of direct messages to transmit

information to the interested users. The reason for this is that `@thespisapp` can be followed by non registered users or people that are interested in the specific application domain but have not yet sign up. Another design principle is that tweets or direct messages mention related users. The rationale is that we try to connect the users and build a user research community around this workflow application.

4 Conclusions

The proposed twitter-enabled workflow creation environment is expected to provide VPH community with a social media enabled environment that facilitates knowledge management and transfer for the VPH community to benefit. Thespis concept allows researchers to design and share custom workflows but also alert users and colleagues of potential new models and workflows of interest. Ratings and usage statistics of existing models or workflows raise community awareness and builds virtual research communities where experts speak with each other and have advise to share. Such tools made available in the open source community are expected to engage more newcomers to the VPH community.

The social networking features of the workflow application presented here are currently available in a prototype version. The application will be evaluated in the context of the "YPERTHEN"[2] and there are still some ideas for future work focusing on providing more personalized information to the users. For example, the application could recommend existing models based on their rating to new workflow users [16] and notify existing modelers about newly inserted models capable of being coupled with existing ones.

There are a number of challenges though. Twitter has recently changed the rules of their API and the new rules restrict the number of tweets or direct messages made per account. This will have an impact to the workflow application when a large number of users need to be notified with direct messages. Another possible limitation is that the content in Twitter is public by default and as is the case with any Web/Cloud based platform the user can not always be pretty sure that his/her data are kept private and used for the intended purpose only. Nevertheless, for this application the data exchange are minimal and only references (through hyperlinks or tags) to the possibly intellectually protected information are shared. In any case, what worries more the expert community is finding ways to ensure that accurate information and validated models are only published online, but nevertheless countering bias and misinformation has been a long lasting concern especially during election periods!

Acknowledgements. This work was supported by the community initiative Program INTERREG III, Project "YPERTHEN", financed by the European Commission through the European Regional Development Fund and by National Funds of Greece and Cyprus, and also the TUMOR (FP7-ICT-2009.5.4-247754) project.

[2] `http://www.yperthen.gr`

References

1. Clapworthy, G., Kohl, P., Gregerson, H., Thomas, S., et al.: Digital human modelling: a global vision and a european perspective. Digital Human Modeling, 549–558 (2007)
2. Marias, K., Dionysiou, D., Sakkalis, V., Graf, N., et al.: Clinically driven design of multi-scale cancer models: the ContraCancrum project paradigm. Interface Focus 1, 450–461 (2011)
3. Hughes, B., Joshi, I., Wareham, J.: Health 2.0 and medicine 2.0: tensions and controversies in the field. Journal of Medical Internet Research 10 (2008)
4. Basdekis, I., Sakkalis, V., Stephanidis, C.: Towards an accessible personal health record. In: Nikita, K.S., Lin, J.C., Fotiadis, D.I., Arredondo Waldmeyer, M.-T. (eds.) MobiHealth 2011. LNICST, vol. 83, pp. 61–68. Springer, Heidelberg (2012)
5. Brownstein, C., Brownstein, J., Williams, D., Wicks, P., Heywood, J.: The power of social networking in medicine. Nature Biotechnology 27, 888–890 (2009)
6. Roniotis, A., Marias, K., Sakkalis, V., Manikis, G.C., Zervakis, M.: Simulating Radiotherapy Effect in High-Grade Glioma by Using Diffusive Modeling and Brain Atlases. Journal of Biomedicine and Biotechnology 2012, 9 (2012)
7. Sakkalis, V., Sfakianakis, S., Marias, K., Stamatakos, G., et al.: The TUMOR Project: Integrating cancer model repositories for supporting predictive oncology. In: 2nd Virtual Physiological Human Conference, VPH 2012 (2012)
8. OReilly, T.: What is web 2.0: Design patterns and business models for the next generation of software. Communications & strategies, 17 (2007)
9. Mandavilli, A.: Peer review: Trial by twitter. Nature 469, 286–287 (2011)
10. Java, A., Song, X., Finin, T., Tseng, B.: Why we twitter: understanding microblogging usage and communities. In: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis, pp. 56–65. ACM (2007)
11. Barker, A., van Hemert, J.: Scientific Workflow: A Survey and Research Directions. In: Wyrzykowski, R., Dongarra, J., Karczewski, K., Wasniewski, J. (eds.) PPAM 2007. LNCS, vol. 4967, pp. 746–753. Springer, Heidelberg (2008)
12. Curcin, V., Ghanem, M.: Scientific workflow systems-can one size fit all? In: Cairo International Biomedical Engineering Conference, pp. 1–9. IEEE (2008)
13. Sfakianakis, S., Sakkalis, V., Marias, K., Stamatakos, G., McKeever, S., Deisboeck, T., Graf, N.: An architecture for integrating cancer model repositories. In: Conf. Proc. IEEE Eng. Med. Biol. Soc., pp. 6828–6831. IEEE (2012)
14. Southern, J., Pitt-Francis, J., Whiteley, J., Stokeley, D., Kobashi, H., Nobes, R., Kadooka, Y., Gavaghan, D., et al.: Multi-scale computational modelling in biology and physiology. Progress in Biophysics and Molecular Biology 96, 60 (2008)
15. Leiba, B.: Oauth web authorization protocol. IEEE Internet Computing 16, 74–77 (2012)
16. Lakiotaki, K., Delias, P., Sakkalis, V., Matsatsinis, N.: User profiling based on multi-criteria analysis: the role of utility functions. Operational Research 9, 3–16 (2009)