QLectives – Socially Intelligent Systems for Quality
Project no. 231200

Instrument: Large-scale integrating project (IP)
Programme: FP7-ICT

Deliverable D.4.2.1

QScience v1 (Web2.0)

Submission date: 2010-03-01

Start date of project: 2009-03-01 Duration: 48 months

Organisation name of lead contractor for this deliverable:
University of Fribourg

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination level		
PU	Public	**X**
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidental, only for members of the consortium (including the Commission Services)	

Document information

1.1 Author(s)

Author	Organisation	E-mail
Matúš Medo	University of Fribourg	matus.medo@unifr.ch

1.2 Other contributors

Name	Organisation	E-mail
Nigel Gilbert	University of Surrey	n.gilbert@surrey.ac.uk
Yanbo Zhou	University of Fribourg	nbboob@gmail.com

1.3 Document history

Version#	Date	Change
V0.1	15 January, 2010	Starting version, template
V0.2	26 January, 2010	First draft completed
V0.3	9 February, 2010	Corrections and stylistic revisions
V1.0	01 March, 2010	Approved version to be submitted to EU

1.4 Document data

Keywords	QScience, Web 2.0, user-provided input, personalized recommendations
Editor address data	matus.medo@unifr.ch
Delivery date	01 March, 10

1.5 Distribution list

Date	Issue	E-mail
	Consortium members	QLECTIVES@list.surrey.ac.uk
	Project officer	Jose.FERNANDEZ-VILLACANAS@ec.europa.eu
	EC archive	INFSO-ICT-231200@ec.europa.eu

QLectives Consortium

This document is part of a research project funded by the ICT Programme of the Commission of the European Communities as grant number ICT-2009-231200.

University of Surrey (Coordinator)
Department of Sociology/Centre
for Research in Social Simulation
Guildford GU2 7XH
Surrey
United Kingdom
Contact person: Prof. Nigel Gilbert
E-mail: n.gilbert@surrey.ac.uk

University of Fribourg
Department of Physics
Fribourg 1700
Switzerland
Contact person: Prof. Yi-Cheng Zhang
E-mail: yi-cheng.zhang@unifr.ch

Technical University of Delft
Department of Software Technology
Delft, 2628 CN
Netherlands
Contact Person: Dr Johan Pouwelse
E-mail: j.a.pouwelse@tudelft.nl

University of Warsaw
Faculty of Psychology
Warsaw 00927
Poland
Contact Person: Prof. Andrzej Nowak
E-mail: nowak@fau.edu

ETH Zurich
Chair of Sociology, in particular
Modelling and Simulation
Zurich, CH-8092
Switzerland
Contact person: Prof. Dirk Helbing
E-mail: dhelbing@ethz.ch

Centre National de la Recherche Scientifique, CNRS
Paris 75006,
France
Contact person: Dr. Camille ROTH
E-mail: camille.roth@polytechnique.edu

University of Szeged
MTA-SZTE Research Group on
Artificial Intelligence
Szeged 6720, Hungary
Contact person: Dr Mark Jelasity
E-mail: jelasity@inf.u-szeged.hu

Institut für Rundfunktechnik GmbH
Munich 80939
Germany
Contact person: Dr. Christoph Dosch
E-mail: dosch@irt.de

QLectives introduction

QLectives is a project bringing together top social modelers, peer-to-peer engineers and physicists to design and deploy next generation self-organising socially intelligent information systems. The project aims to combine three recent trends within information systems:

- **Social networks** - in which people link to others over the Internet to gain value and facilitate collaboration
- **Peer production** - in which people collectively produce informational products and experiences without traditional hierarchies or market incentives
- **Peer-to-Peer systems** - in which software clients running on user machines distribute media and other information without a central server or administrative control

QLectives aims to bring these together to form Quality Collectives, i.e. functional decentralised communities that self-organise and self-maintain for the benefit of the people who comprise them. We aim to generate theory at the social level, design algorithms and deploy prototypes targeted towards two application domains:

- **QMedia** - an interactive peer-to-peer media distribution system (including live streaming), providing fully distributed social filtering and recommendation for quality
- **QScience** - a distributed platform for scientists allowing them to locate or form new communities and quality reviewing mechanisms, which are transparent and promote

The approach of the QLectives project is unique in that it brings together a highly inter-disciplinary team applied to specific real world problems. The project applies a scientific approach to research by formulating theories, applying them to real systems and then performing detailed measurements of system and user behaviour to validate or modify our theories if necessary. The two applications will be based on two existing user communities comprising several thousand people - so-called ”Living labs”, media sharing community tribler.org; and the scientific collaboration forum EconoPhysics.

Executive summary

In 1998, Fribourg university launched the web-based publication "Econophysics Forum" (EF, see `www.unifr.ch/econophysics`) which has served as a community ideas exchange platform for the last ten years. Today, it is a leading web site frequently visited and gathering online publications for the interdisciplinary and complex sciences, counting over 2 000 registered users. Because it was only maintained, not upgraded, the web site gradually became outdated according to current IT standards. This concerned not only the page technology but also the page spirit: the opportunity for user contributions was practically non-existent. For an ordinary user, the only way how to influence the page's content was to send an e-mail to Fribourg-based administrators of the EF.

The Econophysics Forum aims to serve as a test laboratory for ideas and technologies developed within the QLectives project. The goals of our efforts are: (i) to turn the EF into a platform to support scientific quality and creativity through proactively linking scientists with shared interests and quality assessments, (ii) support rapid formation of new scientific communities with rapid consensus on quality and significance of publications and other scientific resources, and (iii) grow social capital through reliable reputation and trust mechanisms and thus improve scientific productivity (share tasks, reduce re-invention).

The first planned step was to make the EF more open to the input from its users and employ some simple recommendation and community-building tools, while still adhering to the standard Web 2.0 approach. This deliverable describes the new func-

tions that have been implemented to attain these ends. These functions can be divided into the following categories:

1. User accounts and basic account administration.
2. Submitting of new resources (news items, information about coming events, book reviews, and scientific papers).
3. Commenting on all resources present in the EF, rating of comments and voting of scientific papers.
4. Information about recently most popular papers and most active users.
5. Personalized recommendation.

Now that the new version of the Econophysics Forum is available, we shall start collecting data with user feedback and behaviour which will be important for the future development of the page.

Contents

Chapter 1

Introduction

In 1998, Fribourg university launched the web-based publication "Econophysics Forum" (EF, see `www.unifr.ch/econophysics`) which has served as a community ideas exchange platform for the last ten years. Today, it is a leading web site frequently visited and gathering online publications for the interdisciplinary and complex sciences, counting over 2 000 registered users. Because it was only maintained, not upgraded, the web site gradually became outdated according to current IT standards. This concerned not only the page technology but also the page spirit: the opportunity for user contributions was practically non-existent. For an ordinary user, the only way how to influence the page's content was to send an e-mail to Fribourg-based administrators of the EF. (That was how informations about future events, book reviews, and links to unnoticed papers made their way to the front page.)

The Econophysics Forum aims to serve as a test laboratory for ideas and technologies developed within the QLectives project. The goals of our efforts are: (i) to turn the EF into a platform to support scientific quality and creativity through proactively linking scientists with shared interests and quality assessments, (ii) support rapid formation of new scientific communities with rapid consensus on quality and significance of publications and other scientific resources, and (iii) grow social capital through reliable reputation and trust mechanisms and thus improve scientific productivity (share

tasks, reduce re-invention).

The first planned step was to make the EF more open to the input from its users and employ some simple recommendation and community-building tools, while still adhering to the standard Web 2.0 approach (see the manifesto "What is Web 2.0" by T. O'Reilly, `http://www.oreilly.com/go/web2`, for detailed discussion of this approach). On the following pages we describe the new functions that we implemented to attain these ends.

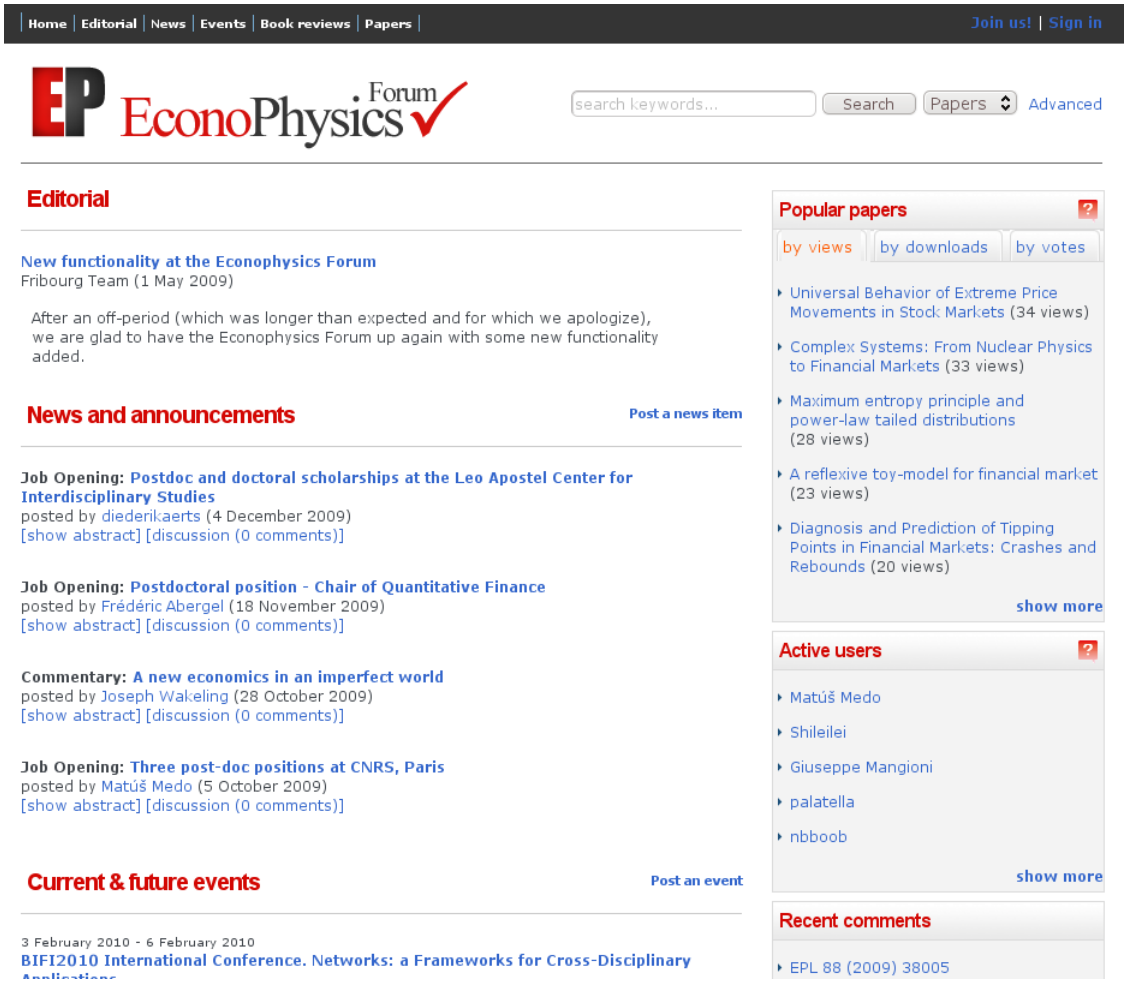

Figure 1. New front page of the Econophysics Forum.

Chapter 2

Description of new functions

Detailed description of the newly implemented functions follows.

2.1 User accounts

The first step to build a web page with trustful user-generated content was to allow users to create accounts. To create an account, only very limited information is needed (see Fig. 1) and hence the whole procedure is easy and fast.

With a user name and password, a user can log in (see Fig. 2) and use all the features available on the web page. Even users without an account (so-called guest users) can see all the content on the page and can contribute new content. The only limitation is that posts by unregistered users, before being displayed on the page, are placed in a queue awaiting editorial approval which can cause a few day's delay in their visibility to other users. In this way, the Econophysics Forum is both secure from uncontrolled content (input by spammers or otherwise malicious users) and at the same time open to input from those who are not willing to spend time by creating a new account. In the current testing phase of the new page design, approximately half of the contributions come from unregistered users, which shows that having a two-fold submission process is useful.

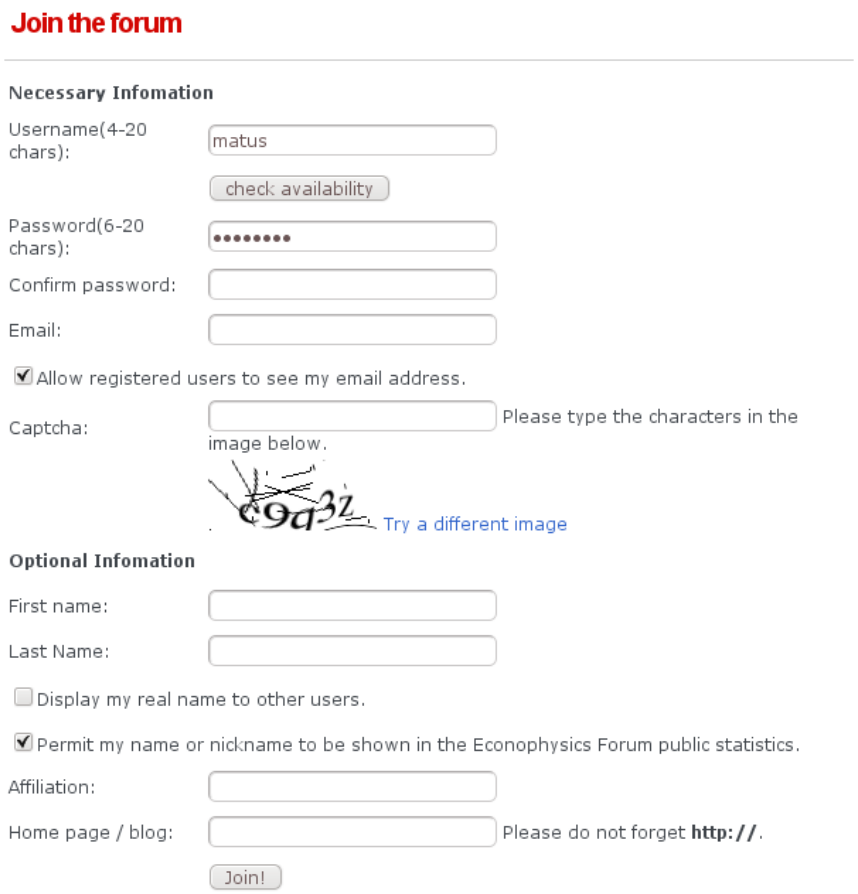

Figure 2. Sign up form for a new user of the EF.

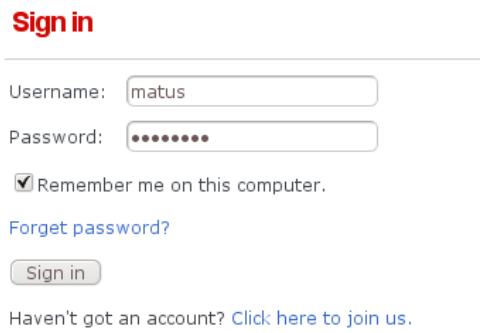

Figure 3. Simple log in form.

2.2 Users' input

As already explained above, allowing for users' input was one of the main goals of the redesign of the EF. Users can now directly contribute to each section except for edito-

rials (where guest editorials are still welcome but before being posted, they first need to be considered by the editors). In a few clicks, any registered user of the EF can post a news item (job opening, link to a relevant popular article or blog, and others), information about a coming event (workshop, conference, etc.), book review, or a scientific paper (see the links "Post a news item" and "Post an event" in Fig. 1).

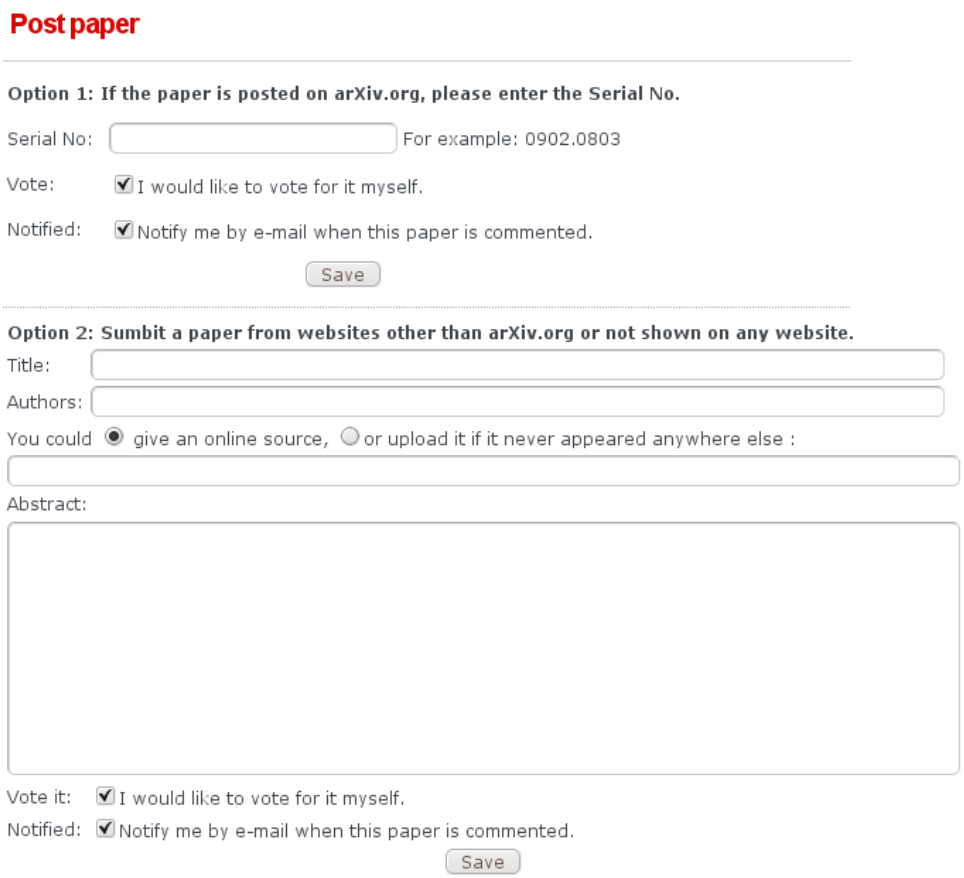

Figure 4. Posting of a new paper.

The Econophysics Forum started as a common place for the latest research in econophysics and interdisciplinary physics. The sharing of relevant scientific papers is still the key feature of this page with the current number of papers in our database exceeding 5 000. There are two different ways to submit a new paper (see Fig. 4). If the paper is available on arXiv (a popular online store for scientific papers available at `arxiv.org`), it is enough to provide the paper's arXiv ID (*e.g.*, 1001.2733) and the

paper's basic bibliographic information (title, list of authors and abstract) is automatically retrieved from arXiv. To submit a paper which is not available on arXiv, the user has to provide information about the paper and either give an online source (an Internet address where the paper is located) or upload the paper directly to the EF (which is useful if the paper has never appeared anywhere else). In both cases, users can vote for the papers they are submitting (see below) and subscribe to notifications about comments to the submitted papers (this is particularly useful to the authors who want to know what others think about their work). For both voting and subscribing to notifications, there are simple check-boxes shown in the submission form (see Fig. 4).

Besides submitting new content, users can also express their opinion about the existing content by comments and votes. Every item on the page (whether it be an editorial or an invitation to a workshop) has its own discussion and anyone can write a comment about it. Further, users can give "votes" to papers to approve their content (one user can vote for one paper at most once). Votes should serve as a better indicator of a paper's quality than quantitative measures (number of views and downloads).

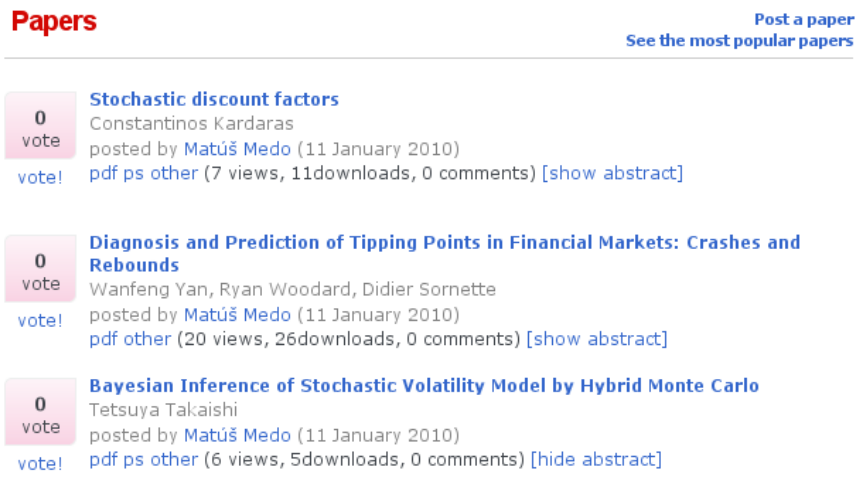

Figure 5. Top of the list of papers.

Finally, to judge the quality of discussion comments (which can be interesting pieces of knowledge themselves but also can be inappropriate, unfair, or even ma-

licious), users can give plus or minus rating to each individual comment (see Fig. 6). This information can help us in future to attach different qualities to each registered user (quality of contributions, quality of comments, etc.).

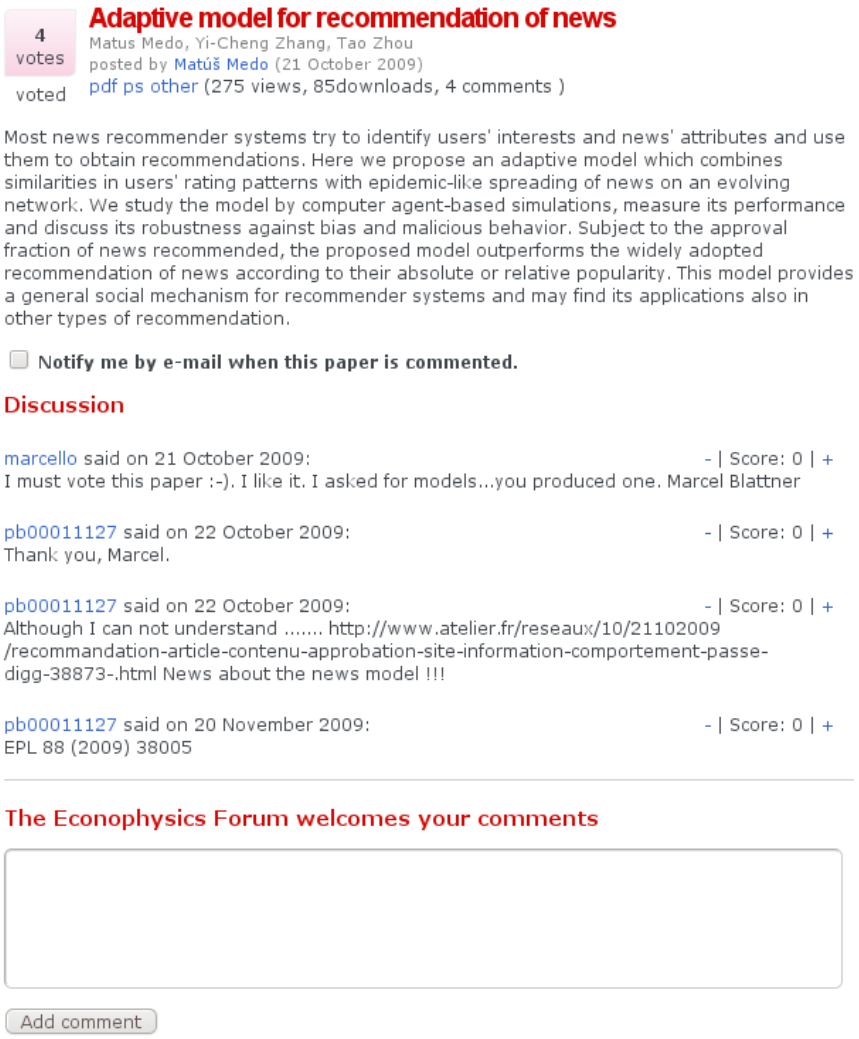

Figure 6. Discussion of a paper.

2.3 Statistics and trends

The new version of the EF traces the behaviour of its users and generates some simple statistics and trends from these data. The results are shown directly on the front page in three distinct information boxes on the right hand side (see Fig. 1) in a similar

way as it is nowadays done on many information portals (compare with, for example, `www.cnn.com`).

Popular papers: The five papers that were most popular in the last week are listed here. Each user can choose whether the ranking should be according to the number of views a paper has, the number of downloads, or the number of votes. Link "show more" points to a longer list of popular papers where, as well the ranking criterion, one can also choose the length of the considered period (see Fig. 7).

Figure 7. List with the most popular papers.

Active users: The five most active users are listed here. The activity of individual users is computed from the number of their paper submissions, comments, and votes during the last week (see Appendix). Link "show more" again points to a more complete list. The overview of most active users could serve in future as an additional motivation for users to contribute.

Recent comments: The last five comments are listed here. This information could help to fuel the discussion by showing the most recent comments to all users of the EF.

In addition to the tools described above, users can subscribe to the RSS feed of the Econophysics Forum. Then they will receive information about all changes on the page to their favourite personal RSS reader (this output, however, is not filtered and hence users receive information about *all* new content on the page).

Papers updated to arXiv are often preprints and hence they may have later improved versions (authors can simply upload them to arXiv where they replace the previous versions). To follow these continuous modifications of papers, we employ a regular check of papers submitted from arXiv to ensure that the title, list of authors, and abstract shown on the EF conform with the latest versions available on arXiv.

2.4 Personal tools

All user names are links pointing to the personal pages of these users. (For example, if user X submitted a paper, his name is shown together with the paper—see Fig. 4—and anyone can find basic information about user X by clicking his name.) This personal page includes the user's e-mail address, link to a homepage, affiliation, and history of the user's contributions (submissions, comments, and votes). Users can opt not to show this information by appropriate settings in their personal profiles.

In contrast to the very basic information available about other users, each user's own personal page includes a personalized recommendation which is not visible to the others (see Fig. 8). This recommendation is two fold. Firstly, there is a list of papers recommended to the given user (this recommendation is obtained on the basis of the user's past activities and a simple paper similarity measure obtained from views, votes, and comments of respective papers). Secondly, users similar to the given user are shown (in this way, one can look for potential collaborators or simply for people to discuss some work with). The list of the user's notifications can be also found here

and can be used for their adjustment. See Appendix for the overview of formulas used for personal recommendation.

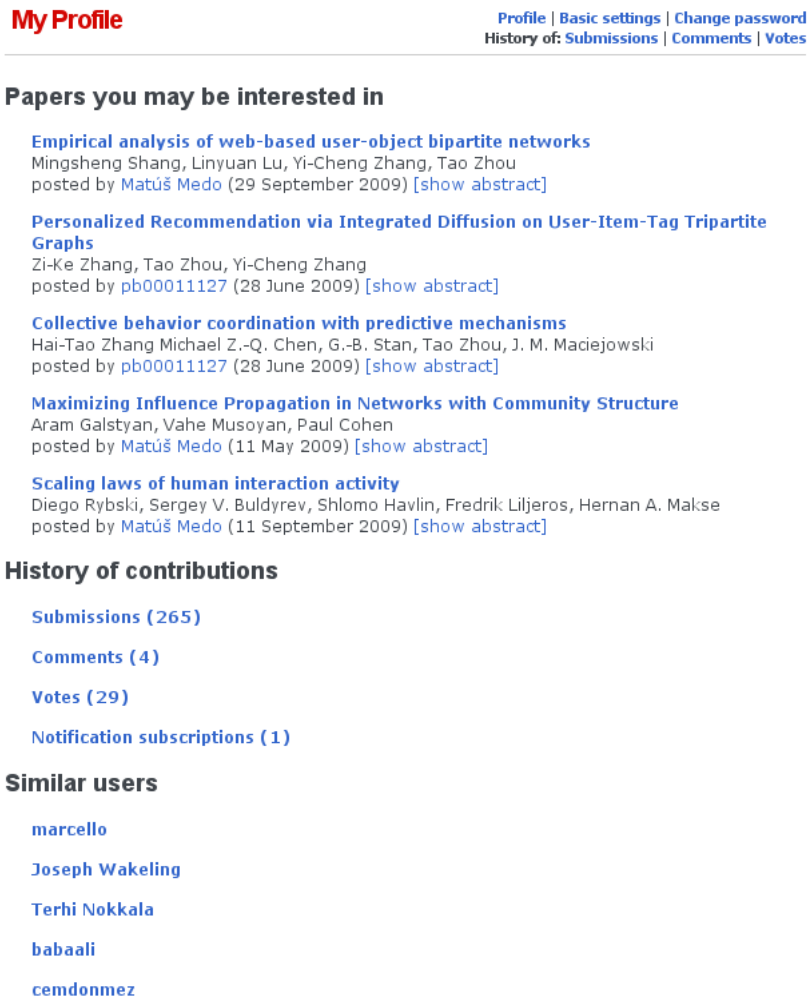

Figure 8. Complete user profile with recommendation.

Chapter 3

Summary and further questions

This deliverable describes the recent major upgrade of the Econophysics Forum, focusing on the new functions that have been introduced to reflect the goals of the QLectives project. We believe that these modifications will stimulate user activity and also transform user behaviour from the passive mode into an active one where users select or create the content together.

In addition to the implemented functions, there is a number of other issues we will target in the near future. Firstly, page design, user interface and user experience need to be studied and optimized. Secondly, used recommendation tools need to be made transparent (such that users understand why particular papers or particular users have been recommended to them—see the recent keynote talk from ACM Recommender Systems 2009, `http://recsys.acm.org/2009/tutorial3.pdf`). Thirdly, additional collaborative tools (such as internal messaging or collaborative file editing) should be added to further support scientific collaboration and creation of research communities.

While the improved page is already available to all users, it is still in a testing phase and has not yet been advertised (out of more than 600 users subscribed to the mailing list of the EF, less than 200 of them currently visit the site regularly). We are looking forward to a public launch of the new functionality as it will provide us not

only with feedback from our users but also with data that will be important for further development of QScience.

Appendix A

Recommendation formulas

Denoting the number of papers submitted by user i as P_i, number of votes given by user i to papers as V_i, and votes given by user i to comments as v_i, activity of user i is computed as

$$A_i = 3P_i + 2V_i + v_i.$$

This particular form reflects the fact that paper submissions are more important than votes. Individual coefficients can be modified to best reflect our priorities and incentives which we want to give to the users.

Interest of user i in paper α, $I_{i\alpha}$, is measured 1 if the user has viewed the paper's abstract, 3 if the user has voted the paper, and 5 if the user has downloaded the paper. If several of these individual actions were made, their contributions are summed together and hence the maximal possible value of interest is $1 + 3 + 5 = 9$.

Similarity of users i and j, s_{ij}, is computed as a scalar product of their interest vectors

$$s_{ij} = \frac{\sum_\alpha I_{i\alpha} I_{j\alpha}}{\left(\sum_\alpha I_{i\alpha}^2\right)^{1/2} \left(\sum_\alpha I_{j\alpha}^2\right)^{1/2}}$$

which is a number between zero (no shared interests) and one (identical interests). For each user, five other users with highest similarity values are shown.

Recommendation score of paper α computed for user i has the form

$$R_{i\alpha} = \sum_{j \neq i} s_{ij} I_{j\alpha}.$$

That means, to user i we recommended objects that attracted big interest from users who are similar to i.