

# Analysis of Portal for Social Network of IT Professionals

*Jiri Kubalik<sup>a</sup>, Kamil Matousek<sup>a</sup>, Jaromir Dolezal<sup>b</sup>, Martin Necasky<sup>c</sup>*

*<sup>a</sup>Department of Cybernetics , Faculty of Electrical Engineering*

*Czech Technical University*

*Technicka 2, 166 27 Prague 6, Czech Republic*

*[kubalik@fel.cvut.cz](mailto:kubalik@fel.cvut.cz), [matousek@fel.cvut.cz](mailto:matousek@fel.cvut.cz)*

*<sup>b</sup>Department of Circuit Theory, Faculty of Electrotechnical Engineering,*

*Czech Technical University, Prague*

*Technicka 2, 166 27 Prague 6, Czech Republic*

*[dolezja7@fel.cvut.cz](mailto:dolezja7@fel.cvut.cz)*

*<sup>c</sup>Department of Software Engineering*

*Faculty of Mathematics and Physics, Charles University in Prague*

*Malostranske nam. 25, 118 00 Prague 1, Czech Republic*

*[necasky@ksi.mff.cuni.cz](mailto:necasky@ksi.mff.cuni.cz)*

**Abstract:** *This paper presents an analysis of selected social networking portals for science, research and innovation as the first step in development of a portal for social network of IT professionals in the Czech Republic. First, actual portal implementations of four portals dedicated to scientific communities of researchers are analyzed, compared from the perspective of provided functions, layout and their interdependence and the main features are summarized. Then, the most important recommendations for the portal being developed are presented. The last part of the paper introduces two concepts of structured profiles - knowledge profile and scientific profile - that are proposed in order to precisely and explicitly express knowledge and scientific alignment of an IT professional. These two profiles will make it possible to search and compare IT specialists, their groups and projects or pair the users to the demands on the market. The paper concludes with a list of typical use cases where the profiles will be utilized.*

**Key words:** Social network, scientific profile, user profile, job portal, semantic search, web portal

## 1. Introduction

The project "Social network of IT professionals" (SoSIReČR - Sociální síť informatiků v regionech ČR) is focused on building a social network of computer scientists in the regions of the Czech Republic<sup>1</sup> to improve the status of computer science in the Czech Republic and its contribution to society and to increase the competitiveness of the Czech Republic in the field of informatics. Among other means used to achieve these objectives, such as conferences, workshops and seminars, or a network of regional contacts, a web portal with social networking features is being developed under the project, which provides an umbrella platform for active cooperation between computer scientists. The portal will serve both as a centralized source of relevant information of various kinds; in addition it will provide services and resources for using this information towards unification of the community of computer scientists in the Czech Republic and more effective communication of current issues and implementation of activities in this area.

The main contribution of the SoSIReČR portal we can see in these areas:

- Facilitating communication between computer scientists within the regions as well as with other entities in the field of education and research, economic and governmental sectors by means of a global forum, and thematically focused forums for centralized discussion on urgent topics in computer science, such as research funding, university curricula, etc.

---

<sup>1</sup> There belong the employees and students of regional colleges and universities and the employees dealing with education, research and development in regional research organizations.

- Provision of access to large amounts of information, materials and references concerning informatics and computer scientists (course materials, documentation of successful projects, etc.).
- Establishing and developing co-operation, creation of job or interest groups, searching for suitable entities (people, institutions, research teams and companies).
- Support of job offers supply and demand.

This paper summarizes the results of the first phase of the portal development, which was a requirements analysis, carried out jointly by the teams of the Czech Technical University in Prague and Charles University in Prague. The first part begins with a brief analysis of four selected international projects aimed at cooperation between scientists, development and education. There are briefly described the basic features of these portals and discussed the most important conclusions resulting from the analysis with regard to the upcoming portal SoSIReČR. The second part will describe the backbone of the social network of computer scientists in the regions of Czech Republic, which shall be provided by the portal. We introduce two important characteristics of persons and other entities in the network: knowledge and scientific profiles. These will enable to describe the knowledge and the scientific focus in a structured way and provide the advanced tools for searching and matching entities in the social network compared with existing solutions. Finally, we demonstrate the advantages of using knowledge and research profiles on basic use cases.

## 2. Analysis of selected portals for science, research and innovation

Abroad, there are similar projects for cooperation between scientists, development and education, that is why we have started first with their analysis. For the analysis, we have chosen the following four portals that are run for at least three years:

- ResearchGate [1],
- SciSpace [2],
- Epernicus [3],
- ScienceStage [4].

We compare the portals in terms of functionalities provided, clarity of their positioning and relationships, and reliability of the portal contents.

### 2.1. ResearchGate

Portal ResearchGate is designed for the cooperation of scientists and sharing of related data. In our opinion, it has the most sophisticated features of the portals analyzed. The portal fully utilizes the options to automatically recommend content, based on user profile which is extensive and well structured. The Portal provides the ability to search for articles in the internal and external databases, not only on the basis of keywords, but also based on an analysis of the freely available article abstracts. At the search results, there is displayed as a percentage the relevance of individual keywords used by the search engine, which acts as a feedback to the portal users, in order to be able to better formulate their questions. Semantic search is also used in the function *Similar Abstract Search*, which allows, on the basis of the article abstract, to find similar articles to the one an author is currently working on. Similarly, the search is being used in the function *Journal Finder*, which, based on the user's article abstract, provides a list of possible article magazines suitable for eventual publication. The portal also offers support for the publication of the offer and demand of jobs.

The main user interface for drawing the verified information is called *ResearchBlog*, which summarizes already published, peer-reviewed articles or important ideas and concepts. For the rapid distribution of information, there is available the *News* information feed and for frequently updated posts one can use *RSS* feeds. Publications at *ResearchBlog* undergo rigorous approval process by the members of the portal. The content of ResearchGate portal can be described as inappropriate, or, conversely, use the "I like this" function for a positive evaluation, which is not anonymous, which contributes to the credibility of content.

The user profile of ResearchGate portal is an example of a good structure, well usable for automatic content recommendation. It contains basic information such as contacts and expertise of the user. In addition, there are a list of user's publications in the internal and external databases (references can be inserted in several formats), a list of favorite literature titles that the user wishes to share with other members, a list of confirmed and verified user's contacts list and user groups of which the user is a member. For publication of a wide range of information there is used a blog. A tool for visualizing

relationships forms another part of the user profile of ResearchGate portal. You can visualize almost all the links of the user profile, such as links to publications by journals and by author's country of origin.

Home page of the portal contains an overview of the portal, email, settings and search. The home page also allows you to customize the portal, there are available e.g. panels *News in My Network*, *Researchers you might know* and *Next Appointments*. The *Researches* part supports direct cooperation between members. Teamwork support is then under *Groups*, where the group's own discussion forum *Discussion board* is available. Functions for content recommendation contribute to support of the work with literature. Extensively used at this portal, there are recommendations for colleague invitations into the system using imports of links (e.g. e-mail contact lists) or the recommendations of a research group.

## 2.2. SciSpace

This portal provides its members only with basic tools of a social network for cooperation. A key aspect of the SciSpace portal can be seen in the complexity of its favorite items *bookmarks*. Bookmarks can be added to the most of portal items. Another function supporting bookmarks, there is a tool called *Bookmarklet* to be integrated into a web browser, which, on pressing the *bookmark it* button, sends a link to be stored in the user's profile at SciSpace portal. For most of the SciSpace portal items, there is also an option to subscribe via RSS. In the SciSpace portal, user groups represent special entities equipped with their own blog, space for file sharing, and all the other items and tools, which are accessible in the profile of an individual user.

## 2.3. Epernicus

Portal Epernicus is more focused on communication and tools for networking of users with a similar focus. User groups include generally available specific communities into which the user must be assigned already at the registration, as well as user-generated optional groups and so called the *Automated networks*, into which a portal member is automatically assigned based on the user profile.

A key aspect of the portal can be seen in the BenchQ application, which provides fast distribution of questions and answers in the portal community. The cooperation efficiency is supported by the possibility to distribute questions using e-mail including an option to reply also using an e-mail message, which is then processed and stored by the portal. The application also enables to specify in detail, what subset of the members of the portal will be asked the question of interest. Notable is also the visualization of links among portal members.

The user profile is comparable by its extent to the ResearchGate portal and it is also well structured for automatic processing. The profile is divided into education and experience, a brief summary of the research focus of the user in the form of a paragraph, specialized knowledge of the user, a list of his/her awards, a list of publications, where presentations and materials from conferences can be inserted along with accompanying information, a list of favorite sites and favorite documents. Moreover, it is possible to insert a file containing user's CV. Users or user profiles, respectively, can be indicated as inappropriate, or, on contrary, they can be given a favorable evaluation called "Kudos" on the Epernicus portal.

## 2.4. ScienceStage

ScienceStage portal is focused primarily on presenting and sharing multimedia for scientific as well as commercial fields. Many well-known foreign universities, such as Yale University, have their websites (*Stages*) here. There are presentations and broadcasts of many television stations and channels of a rather minor character, focused on science. The available tools for working with content correspond to the multimedia character of the portal.

The basic presentation is available on the portal for free, but one can also buy professional variants of a presentation, which include exposure in the front, increased limits of data storage, multimedia player, HD video and other benefits. The portal content can be browsed without registering and logging. A visitor is only limited to content that is indicated as public by the members. Multimedia resources are evaluated by the user on a scale of 1 to 5.

The home page of the logged on user is the best arranged of all the analyzed sites. It perfectly fulfills the function of navigation guide, as it includes links to all portal functions.

In case of the ScienceStage portal, thanks to the portal focus on the multimedia presentation, the user profile is of the same form as the *Stage* presentation. It is divided into only three parts: the basic data,

the address and scientific focus which is selected from the offered list, and a section containing information on work experience, projects, publications and grants. This additional information is not further structured; it is inserted as text paragraphs, which is not very suitable for automatic content recommendation. Privacy settings are available only at favorites. At the *Stage*, in the Blogstream section, comments can be added. The main feature here is the possibility of inserting audiovisual materials, and text documents.

## 2.5. Comparison of Selected Portals

Following are the most important common features of the analyzed social network portals.

*User profile* contains the key information for identification of a user – a member of a portal. Reliable user's identification allows assessing credibility and relevance of the information he/she presents. Based on the user's profile relevant information can be offered to the user as an alternative to classical (passive) searching for the information. Note that the user profile should be well-structured in order to facilitate automated processing and recommending of the portal content.

*Home page of logged in user*, a member of the portal, is the most frequently visited page. This page does not have to be only a plain collection of tabs, links and static data, it can also show adverts to interesting functions the user is not aware of yet. Some portals list the recently invoked functions (or events) here, while others offer a selection of recommended functions based on the user's profile. Such recommendations help to greatly improve usability and intuitiveness of the systems. The goal of these recommendations is to attract the user's attention to respective information channels. Some portals allow personalization of the home page by adding panels of a widget type or others.

*The main portal page* is important for attracting prospective interested users, who are not logged in. Thus, it should offer certain selection of functions to the unregistered visitors as well. For example, the portal ResearchGate allows the users that are not logged in to display job opportunities, basic information about the portal members, and research groups. Detailed profiles and further functions are available only after the registration. The Epernicus portal does not offer any functions to unregistered users, it only shows the public profile of its users. The ScienceStage portal, on the contrary, makes all of the information publically available, no registration is required. Another strategy adopted by some portals is to start running the system as fully open to public for some period of time and then to restrict availability of some functions to registered users only. A system whose functionality is completely restricted only to the logged in users can not attract new users. On the other hand, a system that allows full unrestricted operation for unregistered users does not motivate the users to register. Thus the suitable strategy is somewhere between these two extremes. It is usually implemented so that several functionalities that demonstrate the capabilities of the system are made available to the unregistered visitors so that they can assess the benefit from using the portal. Another way is to let the system fully open for all users while permanently prompting the unregistered users to register to the system.

When comparing the analyzed portals we will focus on the user profiles used and corresponding tools (see Tab. 1) and other selected aspects (see Tab. 2).

**Tab. 1: Detailed comparison of the user profiles**

Elements and functions of user profile	ResearchGate	SciSpace	ScienceStage*	Epernicus
Possibility to close the account	O	O	X	X
Privacy settings	XXX	XXX	X	XX
Favorite bookmarks	X	XXX	X	O
Branch (science)	XX	O	X	X
Blog and discussion	XX	X	X	O
Job related information	X	XX	X	XX
Research related information	XXX	O	X	XXX
Education	XXX	X	X	XXX
Status of the user	O	O	X	O
Profile export	X	O	O	O
Scientific publications	XX	O	X	X
Presentations	X	O	XXX	X

\* Basic unpaid/free version of the user profile is considered (*Stage*).

The symbol “O” in Tab. 1 indicates that the corresponding aspect is not present at the portal at all or only indirectly. For example, “Scientific publications” are not explicitly present at the SciSpace portal; however it is possible to insert an arbitrary file. The number of “X” symbols indicates the extent to which the given aspect is elaborated in the particular portal.

**Tab. 2: Comparison of portals**

Aspects	ResearchGate	SciSpace	ScienceStage	Epernicus
Friends, contacts and invitations	XXX	X	XX	XX
Links to other social networks	XX	O	XX	XX
Groups and support of cooperation	XXX	XXX	XX	XX
Literature, publications and search	XXX	X	X	XX
Notifications	XXX	XX	X	XX
Content reliability and evaluation	XXX	X	XX	XX
Content security	XX	XX	X	X

*Friends, contacts and invitations* represents tools for recruiting new members to the portal such as invitations, recommendations of the portal content to the members as well as potential interested users, support for creating new contacts and importing contacts existing elsewhere, and support for working in groups.

*Links to other social networks* covers means of sharing materials, groups and user’s status at various social networks such as Facebook and others.

*Groups and support of cooperation* covers all tools for cooperation within teams such as tools for creating research, working and interest groups, automated content recommendation as well as project meeting scheduling and management tools.

*Literature, publications and search* represents tools for accessing and searching in various internal and external scientific databases.

*Notifications* covers all means of notifying events such as email, internal messaging, RSS and News channels.

*Content reliability and evaluation* considers content added by members of the portal and refers to means for discussing and evaluating the content.

*Content security* refers to setting the privacy rights to the portal content, groups and other parts except the user profile (this is already considered in Tab. 1).

### 3. Recommendations for SoSIReČR Portal Development

Important features, functions and services that the portal SoSIReČR should provide are:

- *Recruitment of new members*: This is an important aspect since the success of the portal depends on the size and quality of the community of its users. Among frequently used functions for recruitment of new members there belong import of contacts of existing users, recommendations of the portal content to potential users, public presentations of portal members and making a portion of the portal content and functions accessible for unregistered visitors.
- *User-portal interactions*: Semantic web technologies can be used in the portal to understand the meaning of its content that can further make it possible to automatically generate recommended material to the users based on their profiles. Other means of the user-portal communication involve e-mail notifications, discussions, sending contributions to blogs and others.
- *Interactions among users*: This functionality can be realized by internal mail, blogs and other means of instant distribution of questions and answers in the portal community, see the application BenchQ at the portal Epernicus for example.
- *Well-arranged graphical user interface*: From the usability point of view the intuitiveness of portal functions and easy user orientation on the portal is crucial. If possible, all pages of the portal should adhere to the same structure. The most important is the home page of the logged in user, the member of the portal, which is the most frequently visited page. It is advisable to have there, besides static links and information, a dynamically changing part with recommended information, news and functions individually selected for each particular user.

- *User profile*: Reliable user identification allows assessing credibility and relevance of the information the user adds to the portal. Based on the user's profile, relevant information can be actively presented to the user as an alternative to the classical (passive) searching for the specific information. In order to make the portal content processing and selecting the relevant data fully automated, the user profiles have to be rich and well-structured.
- It is not necessary to completely fill in such detailed user profile at the moment of registering to the portal. Instead, only mandatory fields can be filled in at the beginning and the user is later requested to add the remaining data. Then, the profile completeness can be indicated in percents and the requests to fill in missing data can be of the following form: "Please, fill in your career information so that you can make use of the full portal functionalities".

#### 4. Structured Profiles and their Exploitation in SoSIReČR

The central part of our portal is a social network of people, their groups and projects. The people in the network are particular computer specialists from academia, government as well as private sector. They are also users of the portal. In the real world, computer specialists are organized into various research or working groups. The groups then cooperate on solving particular projects. The social network of computer specialists has a complex structure. The aim of the portal is not only to capture the structure but also allow for effective searching in the structure. For this, it is necessary to identify particular characteristics of particular entities in the network.

The portal allows for each specialist, group and project to publish a profile which provides basic information and, also, social links to other specialists, groups and projects. E.g. a person profile contains "classical" information on person's name and employer, group memberships or solved projects. However, we also identified two other very important characteristics during the analysis of the requirements of our users. We call them *knowledge profile* and *scientific profile*. Both are then called *structured profiles* in general. The main difference in comparison to the personal profiles well-known from "classical" social network portals, e.g. Facebook or LinkedIn, is that their precisely and explicitly express knowledge or scientific alignment of a computer specialist, respectively. This allows for precise searching and comparison of computer specialists, their groups and projects or their pairing to the demands on the market.

In the rest of this section, we describe both knowledge and scientific profiles in a more detail. And, we also demonstrate their usefulness on a set of typical use cases which are covered by our portal.

##### 4.1. Scientific profile

Scientific profile is associated with a particular entity, e.g. computer scientist, group or project and specifies its scientific alignment. If associated with a scientist, it characterizes his or her areas of research (s)he is working on. If associated with a research group, it characterizes the research areas of the whole group. The group scientific profile does not necessarily need to be an aggregation of the profiles of the group members. Similarly, if associated with a project, it characterizes the research areas the project fits into.

Formally, a scientific profile is a pair  $S = (A, R)$  where

- $A$  is a set of axes each representing a particular area of computer science, e.g. *databases*, *operating systems* or *software engineering*. Each axis may be further structured into more detailed axes representing sub-areas, e.g. the *operating systems* axis may be further structured into *performance* or *file systems* sub-axes, and
- $R$  is a partial function called *rating* which assigns a number from  $\{0, \dots, 5\}$  to axes or sub-axes in  $A$ .

For a given axis  $a$  from  $A$ , we call  $R(a)$  *rating of a in S*. It expresses the quality of research in the respective (sub-)area as follows:

0. No relationship to the (sub-)area
1. Basic knowledge, interested in the (sub-)area
2. Exploitation of results in own research but no research contributions in the (sub-)area
3. Preliminary research contributions in the (sub-)area
4. Continuously publishing research contributions in the (sub-)area at conferences and journals
5. Research guru in the (sub-)area, publishing own research contributions in top conferences and journals recognized in the (sub-)area.

Rating is a partial function. From the practical point of view, it means that it is not necessary to specify the rating of all axes and sub-axes in the profile. Having two scientific profile, we are able to compare them even their rating is not total. There is also another practical concern. To be able to compare scientific profiles universally, it is also necessary to unify the set of rated axes  $R$ . In other words, it is necessary that each scientific profile processed by the portal uses the same set of axes and sub-axes. Currently, we consider the ACM classification [5] which already defines a sufficient set of axes and is widely applied by scientists to classify their scientific papers.

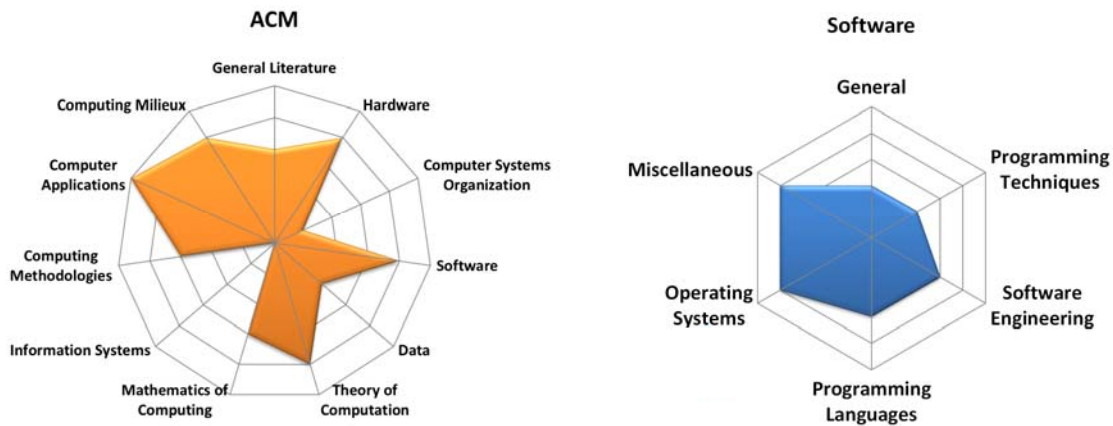


Fig. 1: Sample Scientific Profile

A sample scientific profile (to be correct, only its part) is depicted in Fig.1. The figure shows that scientific profiles may be visualized in a user-friendly way. The left-hand side picture depicts the part of the rating function which assigns a rating to each top-level ACM classification axis. The right-hand side picture then provides a detail of the rating of the sub-axes of the *Software* top-level axis. It is shown only when a user viewing the profile is interested in this particular top-level axis. Suppose that the profile is associated with a computer scientist. Then it specifies that, among other research areas, the scientist is experienced in *Software*. In particular, *operating systems* is one of her main research sub-areas in this area.

#### 4.2. Knowledge profile

Knowledge profiles are very similar to scientific profiles. Formally, a knowledge profile is, again, a pair  $K = (A, R)$  where  $A$  is a set of axes and  $R$  is a rating function. In this case, the axes do not characterize research areas but certain knowledge areas which are important for computer specialists. These include not only knowledge areas typical for computer specialists but also “lighter” areas such as knowledge of legislation, team management or experience in a particular industry, e.g. medicine or automotive industry.

A knowledge profile may be used to characterize knowledge of a particular computer expert. On the other hand, it may characterize a call for a particular computer expert. Or, it may even characterize a curriculum educating students in particular computer science areas.

Again, it is necessary to unify the set of axes  $R$  for knowledge profiles processed by the portal. We employ the axes proposed in [6]. They introduced 17 axes. The rating function assigns, again, each axis with a number from  $\{0, \dots, 5\}$ . Currently, it seems that these axes will not be sufficient for the purposes of the portal. We will probably extend these basic axes with their more detailed classification to sub-axes similarly to the scientific profiles.

#### 4.3. Use Cases

**Searching.** The basic use case is searching for knowledge and scientific profiles which are similar to a given profile. A user fills in a required profile or he selects a predefined profile (e.g. a typical profile of a database architect in Brno). The portal then searches for  $k$  profiles which are the most similar to the given one. Or it can also search for an arbitrary number of profiles with their similarity to the given one not lower than a given  $p$ . This basic usage scenario could be profitably applied to increase awareness of private companies about research by academic teams and, vice versa awareness of academia about the research interests of private companies.

**Comparing.** Profile comparison is another basic use case related to searching. By comparison of two profiles we mean a visualization which clearly shows the differences between both profiles. This may be exploited, e.g. when a user traverses profiles of his or her friends. He/she might see, e.g. how his or her friends, who graduated from the same school, extended their knowledge in particular areas in comparison to his or her own profile. This could be profitably used to motivate computer experts to gradual self-education. If an experts sees how knowledge profiles of his or her friends are being continuously extended, he/she will be forced to complement the missing knowledge.

**Advertising.** The rating function of knowledge and scientific profiles allows for precise specification of "supply and demand" in the social network of computer specialists. For example, a company can precisely specify an ideal candidate on a job position in a form of a knowledge profile of that position. Similarly, a scientist searching for a specific research group with a strong emphasis on a particular research area may advertise this demand by a scientific profile. The portal then may very precisely pair the demands with knowledge and scientific profiles of computer experts, groups and projects in the network.

**Reporting.** Knowledge and scientific profiles allow for creating various statistical reports. For example, it is possible to analyze trends in the evolution of knowledge of computer experts in a particular region. Or, it is possible to watch how the demands of private companies for particular knowledge areas evolve in time.

## References

- [1] Portal of the project ResearchGate, <http://www.researchgate.net/>
- [2] Portal of the project SciSpace, <http://www.scispace.com/>
- [3] Portal of the project Epernicus, <http://www.epernicus.com/>
- [4] Portal of the project ScienceStage, <http://sciencestage.com/>
- [5] *1998 ACM Computing Classification System* [online] on <<http://www.acm.org/about/class/1998/>>.
- [6] VORISEK, J., DOUCEK, P., NOVOTNY, O., *Konkurenceschopnost absolventu IT oboru VS a VOS na trhu prace v CR. Hlavni vysledky projektu.* 15.5.2007. Vysoka skola ekonomicka v Praze. [online] <[http://www.vse.cz/media/konkurenceschopnost\\_it.pdf](http://www.vse.cz/media/konkurenceschopnost_it.pdf)>