

# Large-scale analysis of self-disclosure patterns among online social networks users: A Russian context

Slava Kisilevich<sup>1</sup>, Chee Siang Ang<sup>2</sup> and Mark Last<sup>3</sup>

<sup>1</sup>Department of Computer and Information Science, University of Konstanz, Konstanz, Germany

<sup>2</sup>School of Engineering and Digital Arts, University of Kent, Kent, United Kingdom

<sup>3</sup>Department of Information Systems Engineering, Ben-Gurion University of the Negev, Beer-Sheva, Israel

**Abstract.** Online Social Network Services (SNS) provide an unprecedented rich source of information about millions of users worldwide. However, most existing studies of this emerging phenomenon are limited to relatively small data samples, with an emphasis on mostly “western” online communities (such as Facebook and MySpace users in Western countries). To understand the cultural characteristics of users of online social networks, this paper explores the behavioral patterns of more than 16 million users of a popular social network in the Russian segment of the Internet, namely, My.Mail.Ru (also known as “My World” or “Moj Mir” in Russian). Our main goal is to study the self-disclosure patterns of the site users as a function of their age and gender. We compare the findings of our analysis to the previous studies on Western users of SNS and discuss the culturally distinctive aspects. Our study highlights some important cultural differences in usage patterns among Russian users, which call for further studies in SNS in various cultural contexts.

**Keywords:** Age differences; Correspondence analysis; Cultural differences; Clustering; Gender differences; Information disclosure; Multidimensional scaling; Self-disclosure; Social networking sites

---

## 1. Introduction

Internet studies on social networks and online communities are receiving increasingly more attention. Some recent publications include a peer coordination system for online communities ([Muhlestein and Lim, 2008](#)), a two-stage strategy for

---

*Received Dec 24, 2010*

*Revised Apr 07, 2011*

*Accepted Jun 11, 2011*

finding communities in large social networks (Narasimhamurthy et al, 2010), and anonymization of sensitive information in published social network data (Ying and Wu, 2010; Zhou and Pei, 2010). Due to the surge in online communication, researchers started to explore online self-disclosure (Joinson, 2001; Tidwell and Walther, 2002; Henderson and Gilding, 2004; Christofides et al, 2009). With the latest emergence of online social network services (SNS) such as Myspace<sup>1</sup> and Facebook<sup>2</sup>, studies have begun to examine trust and privacy in this new cyber space (Nosko et al, 2009; Fogel and Nehmad, 2009).

In recent years, these social networking sites have gained increasing popularity among Internet users as they allow for both maintaining relationships with offline friends and forming connections with people one would not have met physically. One important feature of SNS is the user profile, where users can write practically anything about themselves, including personal and sensitive information. In the age of social media, textual information such as name, address and phone number is just the basic. These online systems make it easy to find someone's picture online, even if it is not intended to be shared publicly. The availability of easy uploads and sophisticated search engines have allowed for quick and relatively easy access to details of one's personal life. Even multimedia data such as images and videos can be easily searchable thanks to the practice of folksonomy or social tagging (Halpin et al, 2007).

This new realm of voluntary online information disclosure about themselves has made people not just findable, but knowable, and social media has "brought the voluntary disclosure of personal information to the mainstream" (Madden et al, 2007). In addition, one's reputation is increasingly defined by the information disseminated by others online. In fact, research studies in this area have been escalating in the past few years. However, they tend to focus on the use of SNS in "western" culture. Various studies have shown that the Internet/online communication is not a culturally neutral place (Singh et al, 2003; Callahan, 2005; Pfeil et al, 2006). Furthermore, most studies on self-disclosure on SNSs seem to focus only on a relatively small sample from a city/local community/university (Fogel and Nehmad, 2009; Nosko et al, 2009), often with an emphasis on college students (Ellison et al, 2007; Fogel and Nehmad, 2009). In addition, to the best of our knowledge, most existing studies have focused on self-reported questionnaire study, with only a few studying profiles and behavioral data, i.e. online log of users' behavior in SNS (Pfeil et al, 2009).

Therefore, the motivations of this paper are twofold: (1) to provide an in-depth understanding of self-disclosure patterns in the context of Russian SNS, and (2) to compare the findings to previous research on "western" users of SNSs.

The present study provides a comprehensive examination of the behavioral patterns of more than 16 million users of a popular social network in the Russian segment of the Internet, namely, My.Mail.Ru (also known as "My World"). We performed an analysis of almost the **whole** population of the users, covering a wide range of demographics and use patterns. We carried out a thorough analysis on the types of personal information the users reveal. Furthermore, we also examined how different types of users (particularly age group and gender) disclose information about themselves online.

Specifically, we attempted to address the following research questions:

<sup>1</sup> <http://myspace.com>

<sup>2</sup> <http://facebook.com>

- What are the general cultural characteristics of My.Mail.Ru in the context of Russian segment of the Internet?
- What are the differences and similarities between age and gender in terms of self-disclosure?
- How are ages associated with different patterns of self-disclosure for both genders?
- How are the Russian SNS users different culturally from the previously studied Western-oriented SNS communities?

The rest of the paper is organized as follows. In Section 2, we review current literature on self-disclosure online, relevant studies in SNS, and highlight the issues of cultural difference on the web in general and SNS in particular. In Section 3, we provide an overview to My.Mail.Ru (“My World”), the Russian online social network, the procedure of data collection and data preprocessing. Then in Section 4, we explain our research methodology. Section 5 presents the findings of our analysis while Section 6 discusses the differences between the Russian SNS community under study and Western-centric SNS communities. We conclude the article by outlining future research directions.

## 2. Related Work

### 2.1. Self-disclosure

Self-disclosure is the communicating previously unknown information about oneself so that it becomes shared knowledge, the “process of making the self known to others” (Jourard and Lasakow, 1958, p. 91). It is also defined as “revealing intimate information about one’s self” (Derlega, 1993).

Information disclosure about oneself in human communication is important for various reasons. Firstly, it is essential in forming and maintaining personal relationships (Rubin and Shenker, 1978). Secondly, people continuously compare their disclosure patterns with their peers in order to validate their opinions, perceptions and values. Thirdly, it has been shown that self-disclosure could affect self-esteem and well-being (Leung, 2002; Daley, 2010).

Various research studies have shown that computer-mediated communication (CMC), such as instant messaging and emails, lead to different behavioral patterns in self-disclosure due to certain unique characteristics of such communication, e.g., reduced non-verbal cues, higher control over time and pace of interaction and anonymity (McKenna and Bargh, 1998; Suler, 2004; Walther, 1996). It has been found that self-disclosure was higher online compared to face-to-face communication (Tidwell and Walther, 2002; Joinson and Paine, 2007), even in a non-anonymous CMC setting (Schouten et al, 2007).

Newer CMC technologies such as social media and social networking sites (SNSs) have altered substantially how personal information is disclosed and shared. Unlike emails or instant messaging, where information is generally shared with the person one is directly interacting with, in the context of social media, information is potentially shared with massive groups of people within a social network space. In addition, users are often encouraged by the service providers to share a considerable range of information with a large number of users. Not surprisingly, security and privacy have become recurring concerns (Acoca, 2008).

## 2.2. Relevant studies of Western SNS communities

Studies around the issues of self-disclosure and privacy in SNS are emerging, for instance Nosko et al (2009) carried out a comprehensive study on self-disclosure on Facebook with 400 random sample of a Canadian Facebook community. They found that those seeking a relationship in this community were at greatest risk of privacy threat, and disclosed the greatest amount of highly sensitive and potentially stigmatizing information. They also found age and gender differences in self-disclosure, as identified by earlier research in self-disclosure in other CMC tools in the Western context ([Dindia and Allen, 1992](#); [Bucur et al, 1999](#); [Goodstein, 2007](#)).

Another study by Fogel and Nehmad (2009) examined 205 college students in the US context to investigate risk taking, trust, and privacy concerns in online social networking. One interesting finding was that different SNSs have different levels of trust. They found that Facebook engenders a greater sense of trust than Myspace. Like other studies, they also found that in the Western context, there were some interesting gender effects.

Frye and Dornisch (2010) studied the relationship between perceived privacy and comfort with self-disclosure among 214 US participants. They reported that participants tended to feel more comfortable disclosing information when they perceived the communication tools as offering a higher level of privacy. When studying 704 US SNS student users, Tufekci (2008) also identified gender differences in information disclosure.

Most studies were carried out in the North America and Western Europe cultures and the majority of them focused only on college students or young users. Furthermore, the samples were generally small and often limited to specific community groups (e.g., university, people living in a specific local community).

## 2.3. Cultural differences in the Web and SNS

Although, the Internet is a global phenomenon, its users and creators have different backgrounds, live in different environments, and belong to different cultures. Early studies on cultural difference in cyberspace found that differences in standards of writing dates, addresses, differences in symbols, metaphors could result in confusion or breakdown in communication ([Stengers et al, 2004](#)). Indeed, studies have shown that cultural differences have an impact on people's behavior when communicating through the Internet ([Cakir et al, 2002](#); [Wilson et al, 2002](#)).

The rise of social media or Web 2.0 further highlights the issues of cultural differences, as the focus of these services are largely user-centered. Newer studies in online cultural differences have focused on relationship or social-centric web services such as dating websites ([Rosen et al, 2008](#)) and online social networking. Social networking sites are pervasive. Facebook for instance has currently more than 500 million active users<sup>3</sup>, showing a substantial growth from 61 millions in 2008. It is worth noting that the popularity of social networking sites is not restricted to Western cultures. One such example is Cyworld, a South Korean social network, which has penetrated more than a third of the country's population ([Schonfeld, 2006](#)). Another SNS, Orkut is highly popular in Brazil and

<sup>3</sup> <http://www.facebook.com/press/info.php?statistics>, August 2010

India. The uptake of SNS in Europe is not homogeneous due to linguistic and cultural differences. This results in local SNSs enjoying a substantial share of the market (Fox and Naidu, 2009).

Due to this development, some research studies focusing on the use of SNSs in different cultures are starting to emerge. Marcus and Krishnamurthi (2009) carried out a cross-cultural analysis of SNSs in Japan, Korea and USA to examine the user-interface components of SNSs. They found that unlike Western-centric SNSs, which encourage disclosure of real names, the most popular SNS in Japan, Mixi, focuses on a design that allows Japanese users to maintain a high level of anonymity. Rosen et al (2010) carried out a survey with 425 young users to examine the occurrence of culturally influenced differences in online behavior in SNSs. Using the Hofstede cultural dimension (Hofstede and Hofstede, 2005), they found that users from a “individualistic” cultural backgrounds tended to have larger networks of friends and a greater proportion of friends whom they never met face-to-face. In addition, compared to users from less “individualistic” backgrounds, they shared more photos in SNSs.

In a study about user motivations in using Facebook, Shi et al (2010) found that satisfaction (determined by the ability to maintain offline contacts, entertainment and information seeking) has a significant impact on the continuous intention to use Facebook among Hong Kong users (Shi et al, 2010). Extending this work, Vasalou et al (2010) studied users’ commitment in Facebook in five cultures, i.e. US, UK, Italy, Greece and France and found that culture does effect the user’s motivation for using Facebook.

Although, some studies in and about non-Western cultures are emerging, the majority of studies pertaining to self-disclosure, privacy and trust on SNSs, focus on Western-centric SNS communities, meaning that currently little is known about these issues in other cultures. A preliminary study conducted by Kisilevich and Mansmann (2010) indicated that Russian SNS users tend to disclose less information and were more concerned about the implication on their privacy.

### 3. Data

#### 3.1. My.Mail.Ru Social Network Overview

My.Mail.Ru (“My World”) was launched in 2007 as an online social network that combines all user content on the Mail.ru email portal. In order to register with My.Mail.Ru, the user is required to first open an email account on Mail.Ru, providing the following mandatory fields: *First Name*, *Last Name*, *Birthday* and *Gender*. My.Mail.Ru allows for creating and joining “societies” (i.e. interest groups), sharing photos, videos, music, managing a list of friends, posting messages on the wall, and playing games. It is one of the most popular general purpose social networks in the Russian segment of the Internet. According to its own statistics<sup>4</sup>, it has more than 40 million registered users. However, this number may be overstated due to the fact that the registration with My.Mail.Ru is performed whenever a user opens a new email account.

The user profile consists of seven sections: *Personal info*, *Type*, *About me*, *Interests*, *Education*, *Career*, *Military Service*. The complete list of attributes is

<sup>4</sup> <http://www.corp.mail.ru/about.html>, August 2010

**Table 1.** My.Mail.Ru profile sections and attributes

Section	Attributes	Example
Personal Info	Nickname*	
	First Name*	
	Last Name*	
	Birthday*	12.23.1980
	Gender*	Female, Male
	Location	(Moscow, Russia)
	Relationship	Not married, Married, Divorced
Type	Body/Figure	Slim, Shapely
	Height (cm)	175
	Weight (kg)	84
	Hair Color	Dark, Bright
	Eyes Color	Brown, Bluish
About me	About me	Short description
	Kids	Living with me, Would like to have
	Household	Live with parents, Rent apartments
	Chronotype	Prefer to be active early or late
	Smoking	Not smoking, Smoking
	Alcohol	Drink rarely, Don't drink
	Languages	Russian
	Religion	Christianity
	Political	(Open answers)
	Views	
Interests	Interests	
	Music	
	Movies	
	TV	(Open answers)
	Literature	
	Idols	
Education	-	(Selection from the list)
Career	-	+
Military	-	manual details
Service		

presented in Table 1. The attributes marked with asterisk (\*) are mandatory fields that must be provided by the user during the registration. However, the user may hide *Birthday* information by changing the privacy setting after the profile has been created. Although, the privacy settings allow the user to adjust the visibility of parts of his/her profile to friends only, these settings are not applicable to the Personal info. Therefore, the attributes under *Personal info* section of Table 1 are visible to all users.

### 3.2. Data Collection

We used a website crawling application developed in-house using C# programming language to extract profile information from “My World”. The data collection was performed during a four month period from February 10th 2010 to June 10th 2010 and was divided into several steps. In the first step, we crawled the list of the so called “societies” (or interest groups) totaling in 583,252 groups. In the second step, we acquired the list of members for every group. We ran the crawler on all groups and collected 17,582,267 profiles. Finally, the list of each

person’s friends was retrieved from his/her profile and their profile information was collected. In the end of June 2010, the crawler retrieved the friend list of all users, reaching 27,429,658 profiles and 269,209,683 user-friend relations. During the data collection process, we noticed that many profiles that were previously retrieved were subsequently marked as *blocked* or *deleted* by the social network provider. Therefore, we re-ran the crawler several times on the whole list of retrieved profiles to filter out those profiles which were invalidated by the service provider. At the time of writing this paper (December, 2010), 2,587,984 of all retrieved profiles were invalidated, leaving us with 24,841,674 valid profiles.

### 3.3. Data Selection

In the data preprocessing step we analyzed the age distribution among the SNS users to filter out infrequent and outlier age groups. We selected 21,111,747 users who specified their age and analyzed their age distribution. Fifty percent of the members of “My World” were between age 19 and 33. The median age was 25. 25% of users were aged below 19 and the remaining 25% were above 33 with the maximum age of 53. Ages above 53 were classified as outliers and were filtered out. In addition, we filtered out people whose age was less than 18 for ethical reasons. After the preprocessing, we were left with 16,614,641 profiles between age 18 and 53, where 8,678,282 (52.23%) users were females and 7,936,359 (47.77%) were males.

## 4. Methodology

As mentioned in Section 1, we are interested in the following types of analysis:

1. General cultural characteristics of “My World” in the context of Russian segment of the Internet.
2. Differences and similarities between age and gender of the members of this social network in terms of self-disclosure.
3. Relationship and degree of association between ages and patterns of self-disclosure.

Rather than analyzing a sample of the “My World” users, we have performed our analysis on the **whole** population of the SNS. Though using the entire population eliminates the possibility of a sampling bias, this cannot be considered a strictly controlled experiment as the distributions of some additional factors may vary across genders and age groups. For example, Figure 1 presents the distributions of one factor (“relationship status”) for each gender. Both distributions follow a similar pattern, although, there are some noticeable differences. For instance, the proportion of men looking for a relationship (with a partner) is higher than the proportion of women. On the other hand, there is a higher proportion of married women than married men. These results point to the need of further analysis of the interacting factors in this dataset, which is beyond the scope of this paper.

In order to answer the first research question, we carried out descriptive statistics to explore the differences in behavior between genders with respect



**Fig. 1.** Distributions of types of relationship for each gender (Females - red solid line, Males - blue dashed line. Types of relationship status: *Looking for a Relationship*, *Not Married*, *Married*, *Divorced*, *Widowed*, *In a Sham Marriage*, *Married but looking for a Relationship*, *Engaged*, *Have a boy/girlfriend*, *Available*.

to the disclosed information. This provided an overview of the self-disclosure pattern of the whole population of “My World”.

It was shown in previous studies (Snell et al, 1988; Dindia and Allen, 1992; Fox et al, 2007; Rosen et al, 2008; Frye and Dornisch, 2010) that age and gender are the two most important predictors of online behavior and information disclosure in the cyberspace. However, in some of the previous studies, age was usually grouped into predefined categories, e.g., teenagers between 13 and 19, users over 60 years of age were grouped as “older users” (see for example Pfeil et al (2009)). We used an alternative approach of analysis where every yearly age (18, 19, etc.) was treated separately. This allowed us to analyze the vast amount of SNS users by identifying fine-grained differences between people of different ages.

To be more specific, the problem of finding differences between people of different age groups was addressed by cluster analysis. In cluster analysis, the age groups that are similar to each other are grouped together forming a cluster. Clustering algorithms can be roughly divided into two categories: those which require knowing the number of clusters in advance and those that do not (Maimon and Rokach, 2005). Since our goal is to investigate the age differences and similarities, we would like the clustering algorithm to estimate the number of groups of similar ages. However, the algorithmic process is usually treated as a black box where the researcher obtains the resulting number of clusters based on the input parameters. This “black box” approach is not favorable in the case of social science analysis where the understanding of the underlying process is paramount to answering hypotheses. Therefore, for the purpose of investigating the age difference, we used Multidimensional Scaling (MDS) (Kruskal and Wish, 1978; Borg and Groenen, 1997), which is a visual approach to investigate multivariate data and is usually used for multidimensional data by transforming the distances between cases into a one-dimensional or two-dimensional graph while preserving the relative distances between cases. MDS allows for observing similarities and dissimilarities between ages in terms of information disclosure using graphical representation from which the researcher can determine how ages are



similar or dissimilar using his/her understanding of the underlying structure of the social network.

In addition to finding age differences, we are interested in finding associations between age/gender and the types of information revealed. For example, it is important to identify what age group is more closely associated with revealing *Interests* or *About me* information. We use Correspondence Analysis for this task (Greenacre, 2007), which is a technique for analyzing the degree of correspondence between rows and columns of a contingency table. The degree of association, or correspondence is inherently related to the notion of independence between rows and columns. The higher association between a row and a column is, the more dependent they are. In our case, the higher dependence between a particular age and an attribute of information disclosure indicates the stronger relationship.

## 5. Results and Findings

### 5.1. Descriptive statistics

We used 27 attributes (including age and gender) of the user profile (Table 1) to analyze the patterns of disclosure of each attribute conditioned on age and gender. *Height* and *Weight* were considered as one attribute, since the users either revealed both variables, or none of them. Two  $36 \times 24$  cross tabulations were prepared for both genders where the columns were the disclosed attributes (except for age and gender) and the rows were ages ranging from 18 to 53. The cells of cross tabulations were the ratio of the amount of revealed information (whether the attribute was revealed by a user) divided by the total number of users of that age and gender group. Figure 2 shows 24 line-graphs that visualize the trends of information revelation by age (x-axis) and gender (red solid lines denote the information disclosure by females, while blue dashed lines show how information is disclosed by males). Visual inspection of these graphs suggests that the trends can be roughly divided into five categories (for clarity, the graphs with similar trends are plotted close to each other accompanied by the corresponding category id):

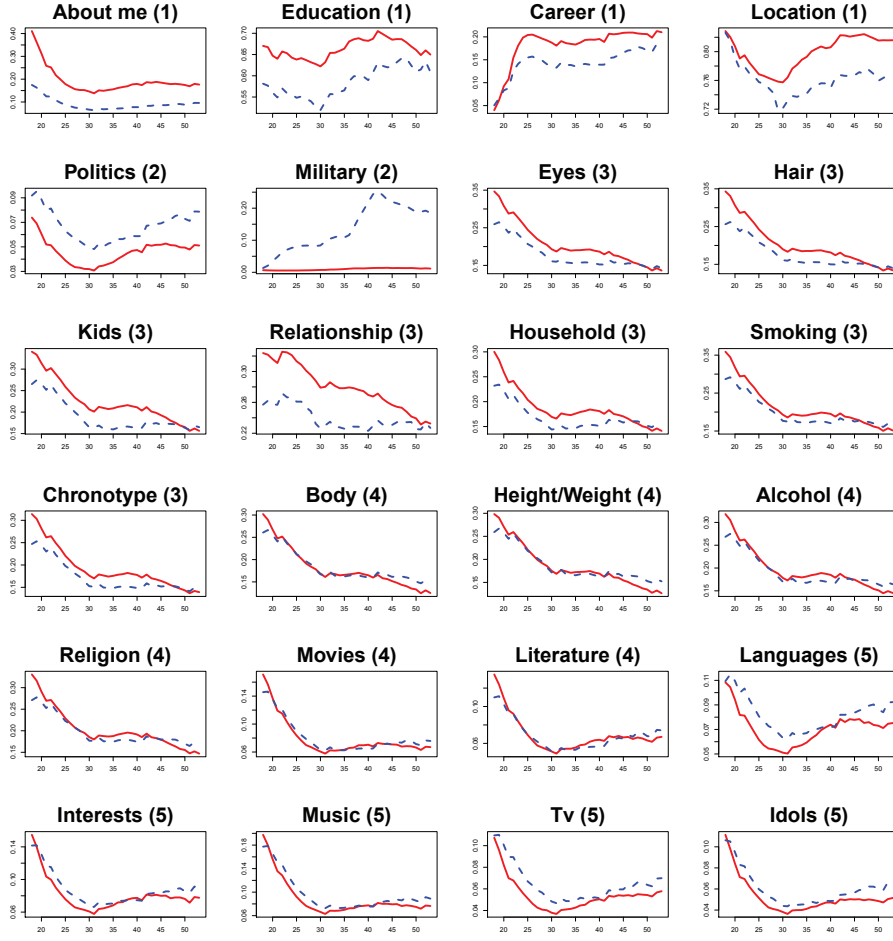
**[Category 1]:** In this category, females of almost all ages reveal substantially more information than males. *About me*, *Education*, *Career* and *Location* belong to this category.

**[Category 2]:** In this category, males of most ages reveal substantially more information than females. *Politics* and *Military* fall into this category.

**[Category 3]:** In this category, females reveal more information almost in every age category but the trend drops around age 45 with cases where males begin to disclose more information around age 53. The following attributes belong to this category: *Eyes*, *Hair*, *Kids*, *Relationship*, *Household*, *Smoking* and *Chronotype*.

**[Category 4]:** This category is very similar to the previous one except for the situation when there is almost an overlap between females and males between ages 20 to 40. Attributes of this category are: *Body*, *Height/Weight*, *Alcohol*, *Religion*, *Movies* and *Literature*.

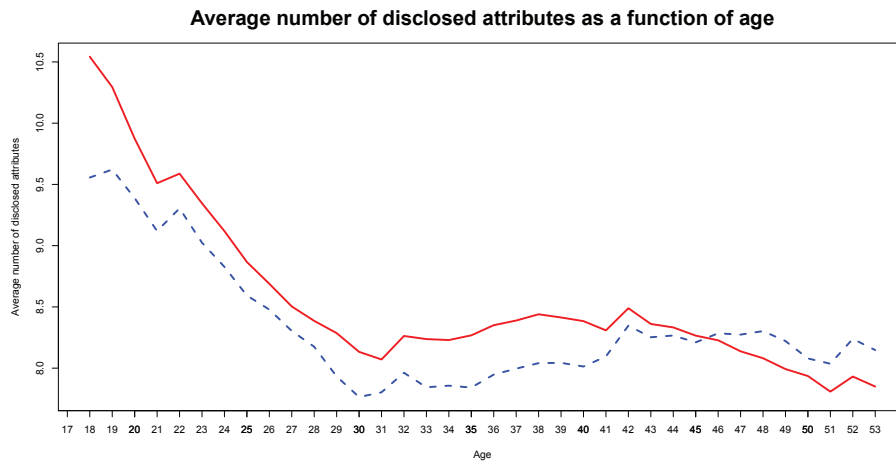
**[Category 5]:** In this category, males reveal slightly more information than females across most of the ages. *Languages*, *Interests*, *Music*, *Tv* and *Idols* are members of this category.



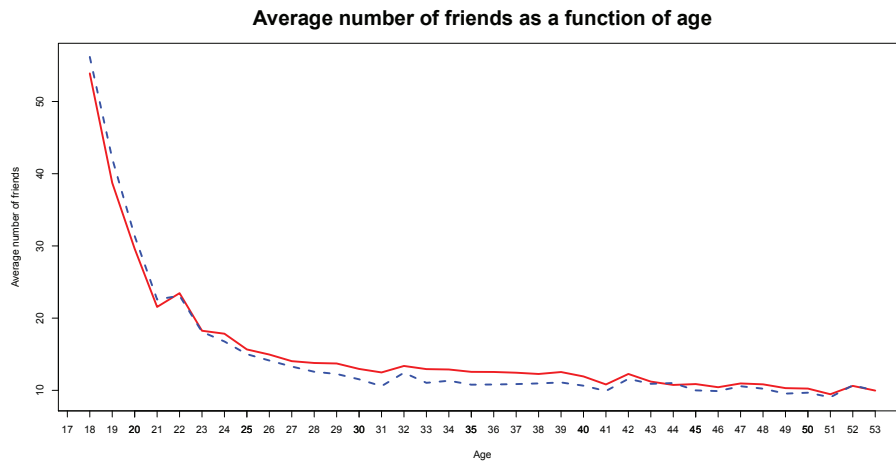
**Fig. 2.** Amount of revealed information by females (red solid lines) and males (blue dashed lines) by age (x-axis). The graphs with similar trends are grouped and labeled by a corresponding number enclosed by parentheses in the title.

In addition, female outperform males in the amount of disclosed information from age 18 in most of the cases except for *Languages*, *Politics*, *Military* and *Career*.

The graph presented in Figure 3 shows the average amount of all revealed attributes conditioned on age and gender. It can be clearly seen that females (red line) tend to disclose more information from the early adulthood to the older ages. This trend continues until age 46 where males (blue dashed line) begin to disclose more information up to age of 53. Figure 4 indicates how the average number of friends changes with age. Males have slightly more friends on average than females between ages 18 to 21. From age 22 up to age 43 and from 45 to 51 females outweigh males in the number of friends, with the exception of age 44 and 52 to 53 where a downturn can be observed.



**Fig. 3.** Average number of disclosed attributes as a function of age and gender. Females are marked by the red solid line, males are denoted by the blue dashed line.



**Fig. 4.** Average number of friends as a function of age and gender. Females are marked by the red solid line, males are denoted by the blue dashed line.

## 5.2. Multidimensional scaling (MDS)

We applied *classical* MDS to the two cross-tabulations, producing two two-dimensional plots with the squared correlation index ( $R^2$ ) of 0.9969 (females) and 0.9984 (males), which is the measure of goodness of fit that reflects the proportion of variance of the distances of the original observations accounted for by the transformed data (the values larger than 0.6 are considered acceptable) and Kruskal's stress type I of 0.032 (females) and 0.025 (males), which according to Kruskal and Wish (1978) are considered as a *good* indication of goodness of fit. Therefore, the two-dimensional solution is adequate. Figure 5 shows a scaled plot of age distances of female members (left) and male members (right).

The left plot on Figure 5 shows two clusters of age groups encircled in green. Ages from 33 to 36 and 36 to 53 appear densely together so we can treat these age groups as similar in terms of information disclosure patterns. All other age groups are located relatively far from each other and it is difficult to outline visually some clusters between ages 18 to 32. However, we can still see relative differences between ages according to the position of age points. For instance, age 19 is more similar to age 18 than to age 20. Age 22 is more similar to age 23 than to age 24.

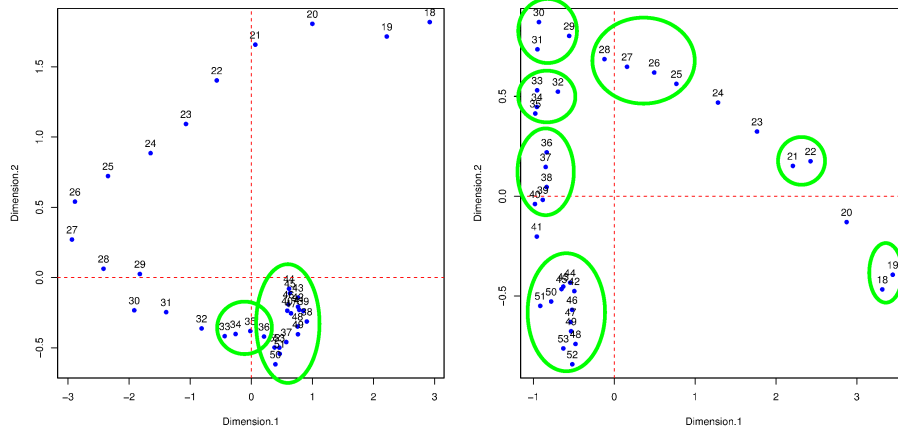
In contrast to females, seven distinguishable age groups can be detected by visual inspection of the MDS plot corresponded to males in Figure 5. The largest groups are 43 to 53, 36 to 40 and 32 to 35, and four groups in which the placement of age points are less dense than in other groups, but still the distance between them is smaller than the distance to other surrounding age points. Such groups are 29 to 31, 25 to 28, 21 to 22 and 18 to 19.

The comparison between MDS plots reveals the difference between genders. The self-disclosure patterns of females seem very heterogeneous until age 32. This is explained by relatively large distances between ages. From age 33 onwards, we observe more similarities between ages, which are characterized by the ability to differentiate two sufficiently large clusters of ages. However, males are more homogeneous in terms of self-disclosure patterns. This allows us to clearly determine seven age groups in which age points are located close to each other suggesting a high similarity between ages.

## 5.3. Correspondence analysis (CA)

Correspondence analysis (CA) is an exploratory technique for analyzing associations between distribution of columns and rows in the contingency table. Like MDS, it aims to reduce the high-dimensional data into a low-dimensional space by plotting transformed high-dimensional vectors of rows or columns on the plot (usually 2D), where the relative position of transformed column or row vectors can be compared for similarity on separate plots. However, CA is a powerful technique for finding associations or relationships between rows and columns. We are interested in the following questions, which can be answered by CA:

- What are the similarities and differences between the ages with respect to the disclosed attributes?
- What is the relationship among the ages and disclosed attributes?



**Fig. 5.** Two-dimensional plots of age categories using classical MDS (Females - left plot, Males - right plot)

- What are the differences between genders with respect to similarities as well as relationships among ages and disclosed attributes?

The initial steps in performing Correspondence Analysis involve:

1. Applying a chi-square test of independence to the contingency table. The test determines whether the rows and columns, or ages and disclosed attributes are independent of each other. The correspondence analysis cannot be performed if columns and rows are independent.
2. Calculation of *row* and *column profiles* that is distribution of information disclosure of each attribute for each age category (row-wise) or distribution of a specific attribute of information disclosure across ages (column-wise).
3. Calculation of *row* and *column masses* - an overall proportion in each column and row.
4. Calculation of *average row profile*, a column-wise proportion of the information revealed in some attribute by all ages.

In CA, the distance between *row* and *column profiles* is measured by chi-squared distance. In the final step, we calculate *inertia*, which is the measure of variance in the multidimensional rows or columns profiles with respect to the centroid calculated as the weighted sum of squared distances from the profiles to their respective centroids. If the two dimensions explain most of the variations, the data can be reduced into two dimensions, which allows to co-plot row and column profiles and to analyze degrees of associations between them. The interpretation of results consists of two parts: (1) interpretation of principal axes, inertia and the contribution of individual profiles to the axes and (2) detection of levels of associations and relationships between rows and columns using graphical representation. The level of association indicates the magnitude of dependency or departure from independence between two categories.

The chi-square test of independence with 805 degrees of freedom and a chi-square value of 894,111 (females) and 904,148 (males) showed at 1% signifi-

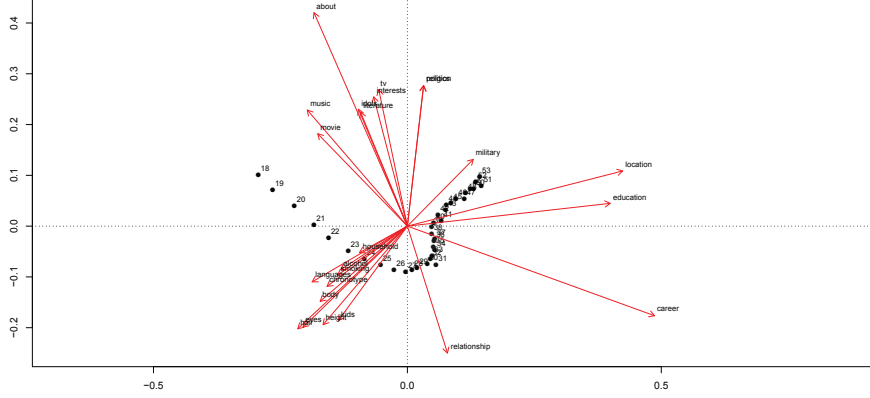
cance level a dependency between ages and disclosed attributes. Therefore, we proceed with the Correspondence Analysis to determine the level of association between ages and disclosed attributes. We applied Correspondence Analysis using R package (Nenadic and Greenacre, 2007) on the two cross-tabulations described in Section 5.1. The cumulative proportion of total inertia explained by two dimensions is 98.1% (females) and 94.9% (males), where the first principal axis ( $\lambda_{1f}=0.0151$ ) accounts for 78.8% (females) and ( $\lambda_{1m}=0.0183$ ) 82.4% (males) of the spatial variation in the data, while the second principal axis ( $\lambda_{2f}=0.0037$ ) accounts for 19.4% (females) and ( $\lambda_{2m}=0.0028$ ) 12.5% (males) of the variation. The total inertia is 0.0191 (females) and 0.0222 (males).

We turn to the analysis of contribution of individual profiles on the axes by analyzing absolute contributions of row and column profiles on the inertia. The absolute contributions indicate the proportion of variance explained by each profile in relation to each principal axis and represent the importance of the profile in determining the direction of the principal axis. The criteria for selection of the profiles that contribute significantly to the principal axis is to find those profiles whose contribution is higher than the average of the whole contributions. Thus, ages 18 to 23 and 49 to 53 (females) and 18 to 24, 40 to 47, and 51 (males) contribute to the first axis. Ages 18 to 23 and 49 to 53 (females) and 18 to 24, 40 to 47, and 51 (males) lie on the opposite sides of the first (vertical) axis (see biplots on Figure 6 for females and Figure 7 for males). Therefore, we can conclude that the first axis opposes young members of the social network to the older group of members. In a similar manner, we perform the analysis of the second (horizontal) axis. Ages 18 to 19, 24 to 33, 48 to 53 (females), and 18 to 20, 25 to 36, 52 to 53 (males) contribute to the second axis. According to the biplots, we conclude that the second axis opposes young members (18-20 for females, 18-22 for males) and very old members of the social network (39-53 for both genders) to the rest of ages.

Likewise, *Eyes, Hair, Location, Education, Career* (females) and *Military, Education, Career* (males) contribute mainly to the first axis, where the first axis in the case of females opposes physical characteristics (*Eyes, Hair*) together with a basic information (*Location*) to the users' records (*Education* and *Career*) and in the case of males the first axis describes users' record information (*Military, Education, Career*).

*About Me, Relations, Religion, Politics, Interests, Music, Tv, Literature, Idols* (females) and *About Me, Relations, Politics, Interests, Music, Tv, Movies, Literature, Location, Military, Career, Idols* (males) contribute mainly to the second axis, where, in the case of females, views (*About Me, Religion, Politics*) and hobbies (*Interests, Music, Tv, Literature, Idols*) oppose the social status (*Relations*) and in the case of males, views (*About Me, Politics*), hobbies (*Interests, Music, Tv, Literature, Idols*) and part of the users' records (*Military*) oppose *Career* and *Location*.

In order to analyze the effect of age on disclosed attributes, we turn to the biplots presented on Figure 6 and Figure 7. The approach we used to draw the biplots, described in detail in Greenacre (2007), is based on the scalar products between row and column vectors and depends on the vector lengths and angles. The biplots were created using age years (rows) expressed in their principal coordinates and the columns in their vertex directions rescaled by multiplying the standard coordinates by the square root of the mass of each attribute of information disclosure as suggested by Greenacre (2007). The key to interpretation of these biplots is to observe the direction of the vector of disclosed attributes

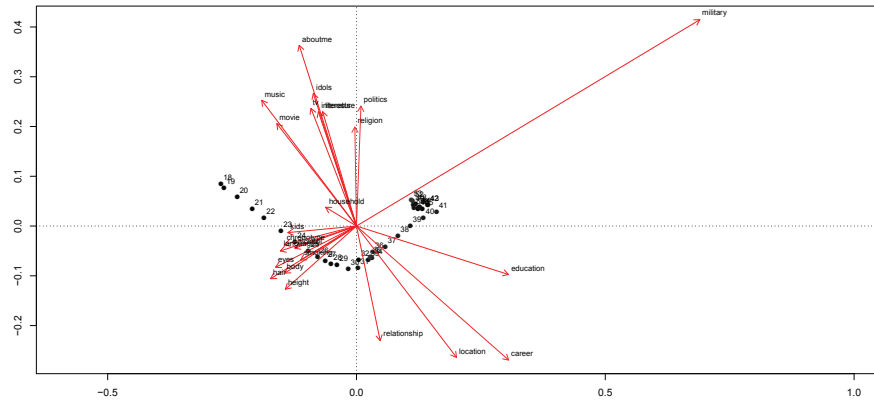


**Fig. 6.** Correspondence Analysis. Biplot of the information disclosure of females by age. Age categories are represented in principal coordinates. Disclosed attributes are represented in their vertex directions rescaled by multiplying the standard coordinates by the square root of the mass of each attribute (column).

(described with arrows on Figure 6 and Figure 7). The projection of an age year onto the vector, which is equal to the length of age vector multiplied by the cosine of the angle between age and attribute of information disclosure, shows the magnitude of the association between age and the particular attribute. The less the angle between the age and the attribute, the higher magnitude the age gets when projected onto the attribute and the more association it has with the attribute of information disclosure.

In case of females (Figure 6), we observe that *Military* is associated strongly with older ages, while younger ages such as 18 to 34 are negatively associated with *Military* (the sign of a scalar product between vectors is negative). Ages 18 to 26 are more than other ages associated with revealing attributes of *Physical* type (*Eyes*, *Hair*, *Height*, *Body*), parts of the attributes of *Social* type (*Kids*, *Household*) and *Chronotype*. Younger people are more associated with *Hobbies* like *Movies*, *Music Idols*, *Literature* than other ages. *Location* and *Education* is connected to ages 31 to 53, where the strongest association with *Location* and *Education* is at the age of 53. Ages 31 to 34 and 47 to 51 are more associated with revealing *Career* than other ages, while ages 26 to 31 are more associated with revealing *Relationship*.

In the case of males (Figure 7), ages 29 to 31 are more associated with revealing *Relationship* than other ages. Ages 30 to 37 are more associated with *Career*, ages 37 to 53 are more associated with *Education* with age 41 having the strongest association, while ages 18 to 22 are more associated with disclosing attributes that belong to *Hobbies* (*Music*, *Movies*, *Tv*, *Idols*, *Literature*, *Interests*). Ages 26 to 53 have almost no association with *Hobbies*, while some have negative associations. *Military* is associated with older ages (37 to 53). We can also observe some association of young members (18 to 19) and old members (40 to 53) with *Religion* and *Politics*, while ages 23 to 39 have almost zero or negative association.



**Fig. 7.** Correspondence Analysis. Biplot of the information disclosure of males by age. Age categories are represented in principal coordinates. Disclosed attributes are represented in their vertex directions rescaled by multiplying the standard coordinates by the square root of the mass of each attribute (column).

## 6. Discussion

The present study showed a number of interesting disclosure patterns, gender differences, age differences as well as associations between ages and disclosure of various types of information. In this section, we discuss how these patterns are different from or similar to previous studies conducted in the Western context.

### 6.1. Gender differences

Indeed, many studies in the Western contexts found that in general females are high disclosers when compared to males in traditional social interaction. However, in an online setting, it was found that self-disclosure was of greater concern to women than men. It was claimed that males were more likely than females to use technology (Vekiri and Chronaki, 2008), and felt more comfortable with technology (Hargittai and Shafer, 2006). This results in men revealing more information online than offline, and could potentially offset the gender differences in offline information disclosure. For instance, a study on Facebook showed no gender differences in information disclosure (Nosko et al, 2009). Although overall the amount of information disclosed by both genders was similar in the Western context, both genders tended to reveal different types of information. Males were generally less willing to disclose their emotions than females, both offline (Snell et al, 1988) and online (Fox et al, 2007). The Tufekci (2008) study of MySpace student users in the US indicated that women were more likely to disclose their interests (including music and books). In the Russian context,



our study showed rather different trends. Russian females were found to have disclosed substantially more information in general across nearly the entire age range, up to 45, from which males started to disclose more.

Unlike results of [Tufekci \(2008\)](#), our study (Figure 2) showed that there was no substantial difference between men and women in disclosing their interests (except for “favorite TV”, in which males disclosed significantly more than females). However, women were more inclined to revealing their physical characteristics such as “eye color” and “hair color”, as well as their social conditions such as relationship status, household, and kids. The only information types men disclosed significantly more were politics, religion views and military records.

It appears that our results are in line with the claim that female interactions are more emotional and feeling-oriented while male’s interactions tend to be instrumental ([Boneva et al, 2001](#)). However, what is slightly surprising is that females seem to reveal more on their education and career in online social networks than men. Our guess is that due to gender inequality in their physical life, Russian women see the online environment as a tool to empower their social status. According to the Russia Longitudinal Monitoring Survey (RLMS) data, majority of men and women unanimously stated that men have better chances at employment. Analyses of Russian job vacancy indicated that up to 30% of job advertisements explicitly stated a desired gender of candidates by the employers, even for professions where gender should not matter. It was also found that Russian women have the tendency of self stereotyping as they believe they are often treated as less preferable employees than men. Another possible explanation lies in finding that women prefer linking up with existing friends in SNSs ([Tufekci, 2008](#)). Since “My World” has a feature of searching for colleagues who (used to) work or study together, female users might have made use of it, thus disclosing more in career and education. In addition, it was also observed that the gender difference gap decreases with age. This could be due to the recent feminization of education in Russian with 57% of women in the higher education ([ROSSTAT, 2004](#)).

Secondly, physical properties are revealed significantly more by Russian women than men. Perhaps physical attractiveness is an important element in their social life. As [Azhgikhina and Coscilo](#) pointed out, the idea of female beauty is highly valued in Russian society ([Azhgikhina and Goscilo, 1996](#)). It is believed that women’s pre-occupation with physical appearance is a struggle against the exhaustion of reality and a means of self expression. Our study showed that even in a text-based communication medium such as this, expression of physical appearance was still an important element for Russian females. Finally, it was also found that in online dating, females preferred males who were low self-disclosers while males showed a slight preference for high self disclosers ([Rosen et al, 2008](#)). Although “My World” is designed to be a generic SNS service, this result may suggest that a substantial number of users were using it as an online dating site. This could explain why female users generally disclosed more than male users.

## 6.2. Age differences

In term of age differences, various studies on adult users in Western SNS have confirmed that younger adults disclosed more personal details online because they may be less cautious and feel more comfortable with online communication ([Goodstein, 2007](#)). Moreover, older people use technology less and thus may not

feel comfortable with online interaction (Bucur et al, 1999). A study on Facebook in a Canadian community found that as age increased, the amount of information presented in personal profile decreased in general. In the context of Russian SNS users, we found similar trends in self-disclosure. Figure 3 shows that in general, for adult users (18 or over) the amount of information decreased as the user got older. In fact, studies in personalities of Russian people indicated that young people (< 35 years) were more extroverted and open than older users, and this finding was cross culturally uniform (Allik et al, 2009).

Our analysis of the whole population of “My World” users, however, revealed a more fine-grained pattern of information disclosure in relation to age. In particular, our analysis showed that although in general, information disclosure decreased with age, the decrease seemed to slow down after the age of 30. It appeared that age 30 was an important turning point for disclosure patterns. Perhaps this is related to the diminishing change in personality after age 30 (McCrae and Costa, 2003), a phenomenon observed in cross cultural context, including Russia (Allik et al, 2009).

Further examining the details of age differences, we observed some associations between certain age ranges and information disclosure attributes from the correspondence analysis. Unsurprisingly, younger users (both male and female) tended to have strong associations with interests related information disclosure such as favorite music and movies. This is similar to the Western-context. For instance Myspace, an SNS which has more younger users, puts a lot of emphasis on music and video.

As aforementioned, female users disclosed significantly more in physical attributes, and correspondence analysis (Figures 6 and 7) showed that younger females between 18 and 26 years old were more strongly associated with physical information disclosure. Furthermore, female users seemed to disclose relationship information at a younger age (26 to 31) than male users (29 to 31).

Although, adult users revealed less information as they grew older, we observed an interesting pattern for users in their early 20s, 30s, 40s and 50s, where there was a “local peak” of information disclosure for these age groups. Perhaps these could be an indication of reaching “significant stages” in life when people from these age groups ventured into new aspects of life (e.g. military service, higher education, employment, etc) thus certain previously irrelevant types of information suddenly became relevant to them.

### 6.3. Trends in number of friends

In the Western context, it was found that men tended to have more “friends” than women as women may care more about whom their “friends” really were and more selective of who they built a relationship with online (Fogel and Nehmad, 2009). In our Russian SNS community, we found the contrary. Here, women users seemed to have slightly more friends on average (mean = 15.366, SD=8.839) than men (mean = 14.795, SD=9.741). We also found that in general, older users have fewer friends than younger users, which is in line with the MySpace study conducted by (Pfeil et al, 2009). This observation has been explained by Boyd (2008) who found that the number of friends teenagers have indicate their “coolness”. Analysis of the whole population in our study shows a subtle yet interesting pattern in the average number of friends across age ranges. Not unlike the pattern of information disclosure (Figure 2), there was also a “local peak” for users in their

early 20s, 30s, 40s and 50s. That implies that as at these ages, a slight increase in information disclosure corresponded to a similar increase in number of friends. In fact, we also found a significant correlation between the total amount of information disclosure and the average number of friends for both female ( $r=.916$ ,  $p=.00$ ) and male ( $r=.834$ ,  $p=.00$ ) users. We believe it is an interesting observation worth further analysis in future studies.

## 7. Conclusions

This paper has provided a comprehensive examination of the behavioral patterns of more than 16 million users of a popular social network in the Russian segment of the Internet, My.Mail.Ru. We performed the analysis collecting data from nearly the entire population of the SNS users, covering a wide range of demographics and usage patterns. We carried out a thorough analysis of the types of personal information revealed by the users and examined how different types of users (e.g. age group, gender) disclose information about themselves online. Throughout this study, we identified various interesting patterns of self-disclosure in both genders across age groups and offered some plausible explanations. In summary, our results suggested that when comparing online information disclosure patterns between Western and Russian SNS users, gender appears to be a culturally distinctive factor while age factor is rather culturally uniform.

We believe that further studies of massive SNS datasets in different countries need to be carried out to understand these interesting issues in more depth. Such studies should also look into additional factors, beyond age and gender, that may affect the online behavioral patterns. The results of such studies can be utilized by the SNS providers worldwide to enter new markets, to improve the user experience, and to better satisfy the privacy needs of specific user groups.

**Acknowledgements.** This work was partially funded by the German Research Society (DFG) under grant GK-1042 (Research Training Group “Explorative Analysis and Visualization of Large Information Spaces”).

## References

- [Acoca B \(2008\) Scoping paper on online identity theft. Tech. rep., Ministerial Background Report DSTI/CP\(2007\)3/Final](#)
- [Allik J, Realo A, Mõttus R, Pullmann H, Trifonova A, McCrae R, et al \(2009\) Personality traits of Russians from the observer’s perspective. \*European Journal of Personality\* 23\(7\):567–588](#)
- Azhgikhina N, Goscilo H (1996) *Getting under their skin! The beauty salon in Russian women’s lives*. Indiana University Press
- Boneva B, Kraut R, Frohlich D (2001) Using e-mail for personal relationships: The difference gender makes. *American Behavioral Scientist* 45(3):530
- [Borg I, Groenen P \(1997\) Modern multidimensional scaling: Theory and applications. Springer Verlag](#)
- Boyd D (2008) Why youth (heart) social network sites: The role of networked publics in teenage social life . In: *Youth, identity, and digital media*, The MIT Press

- Bucur A, Renold C, Henke M (1999) How do older netcitizens compare with their younger counterparts? *CyberPsychology & Behavior* 2(6):505–513
- Cakir H, Bichelmeyer B, Cagiltay K (2002) Effects of cultural differences on e-mail communication in multicultural environments. In: Proceedings of the 3rd international conference on cultural attitudes towards technology and communication, pp 29–50
- Callahan E (2005) Cultural similarities and differences in the design of university Web sites. *Computer-Mediated Communication* 11(1):239–273
- Christofides E, Muise A, Desmarais S (2009) Information Disclosure and Control on Facebook: Are They Two Sides of the Same Coin or Two Different Processes? *Cyberpsychology & Behavior* 12(3):341–345
- Daley A (2010) Being Recognized, Accepted, and Affirmed: Self-Disclosure of Lesbian/Queer Sexuality Within Psychiatric and Mental Health Service Settings. *Social Work in Mental Health* 8(4):336–355
- Derlega V (1993) *Self-disclosure*. Sage Publications
- Dindia K, Allen M (1992) Sex differences in self-disclosure: A meta-analysis. *Psychological Bulletin* 112(1):106–124
- Ellison N, Steinfield C, Lampe C (2007) The benefits of Facebook “friends:” social capital and college students’ use of online social network sites. *Computer-Mediated Communication* 12(4):1143–1168
- Fogel J, Nehmad E (2009) Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in Human Behavior* 25(1):153–160
- Fox A, Bukatko D, Hallahan M, Crawford M (2007) The medium makes a difference: Gender similarities and differences in instant messaging. *Language and Social Psychology* 26(4):389
- Fox D, Naidu S (2009) Usability evaluation of three social networking sites. *Usability News* 11(1):1–11
- Frye N, Dornisch M (2010) When is trust not enough? The role of perceived privacy of communication tools in comfort with self-disclosure. *Computers in Human Behavior* 25(5):1120–1127
- Goodstein A (2007) *Totally wired: what teens and tweens are really doing online*. St. Martin’s Griffin
- Greenacre M (2007) *Correspondence analysis in practice*. Chapman & Hall
- Halpin H, Robu V, Shepherd H (2007) The complex dynamics of collaborative tagging. In: Proceedings of the 16th international conference on World Wide Web, ACM, pp 211 – 220
- Hargittai E, Shafer S (2006) Differences in Actual and Perceived Online Skills: The Role of Gender\*. *Social Science Quarterly* 87(2):432–448
- Henderson S, Gilding M (2004) ‘I’ve never clicked this much with anyone in my life’: trust and hyperpersonal communication in online friendships. *New Media & Society* 6(4):487–506
- Hofstede G, Hofstede G (2005) *Cultures and organizations: Software of the mind*. McGraw-Hill London
- Joinson A (2001) Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology* 31(2):177–192
- Joinson A, Paine C (2007) Self-disclosure, privacy and the Internet. *Oxford handbook of Internet psychology* pp 237–52

- Jourard S, Lasakow P (1958) Some factors in self-disclosure. *Abnormal and Social Psychology* 56(1):91–98
- Kisilevich S, Mansmann F (2010) Analysis of privacy in online social networks of Runet. In: *Proceedings of the 3rd International Conference on Security of Information and Networks*, ACM, pp 46–55
- Kruskal J, Wish M (1978) *Multidimensional scaling*. Sage Publications, Inc
- Leung L (2002) Loneliness, self-disclosure, and ICQ (“I Seek You”) use. *CyberPsychology & Behavior* 5(3):241–251
- Madden M, Fox S, Smith A, Vitak J (2007) *Digital Footprints: Online identity management and search in the age of transparency*. Pew Internet & American Life Project
- Maimon O, Rokach L (2005) *Data mining and knowledge discovery handbook*. Springer-Verlag New York Inc
- Marcus A, Krishnamurthi N (2009) Cross-Cultural Analysis of Social Network Services in Japan, Korea, and the USA. *Internationalization, Design and Global Development* 5623:59–68
- McCrae R, Costa P (2003) *Personality in adulthood: A five-factor theory perspective*. The Guilford Press
- McKenna K, Bargh J (1998) Coming out in the age of the Internet: Identity “demarginalization” through virtual group participation. *Personality and Social Psychology* 75(3):681–694
- Muhlestein D, Lim S (2008) Online learning with social computing based interest sharing. *Knowledge and Information Systems* pp 1–28
- Narasimhamurthy A, Greene D, Hurley N, Cunningham P (2010) Partitioning large networks without breaking communities. *Knowledge and Information Systems* pp 1–25
- Nenadic O, Greenacre M (2007) Correspondence analysis in R, with two- and three-dimensional graphics: The ca palackage. *Statistical Software*
- Nosko A, Wood E, Molema S (2009) All about me: Disclosure in online social networking profiles: The case of FACEBOOK. *Computers in Human Behavior* 26(3):406–418
- Pfeil U, Zaphiris P, Ang C (2006) Cultural differences in collaborative authoring of Wikipedia. *Computer-Mediated Communication* 12(1):88–113
- Pfeil U, Arjan R, Zaphiris P (2009) Age differences in online social networking-A study of user profiles and the social capital divide among teenagers and older users in MySpace. *Computers in Human Behavior* 25(3):643–654
- Rosen D, Stefanone MA, Lackaff D (2010) Online and Offline Social Networks: Investigating Culturally-Specific Behavior and Satisfaction. In: *43rd Hawaii International Conference on System Sciences*
- Rosen LD, Cheever NA, Cummings C, Felt J (2008) The impact of emotionality and self-disclosure on online dating versus traditional dating. *Computers in Human Behavior* 24(5):2124–2157
- ROSSTAT (2004) Women and Men of Russia. *Statistical Digest* p 79
- Rubin Z, Shenker S (1978) Friendship, proximity, and self-disclosure. *Personality* 46(1):1–22
- Schonfeld E (2006) Cyworld attacks! *Business* 20 7(7):84–89
- Schouten A, Valkenburg P, Peter J (2007) *Precursors and Underlying Processes*

- of Adolescents' Online Self-Disclosure: Developing and Testing an Internet-Attribute-Perception Model. *Media Psychology* 10(2):292–315
- [Shi N, Lee M, Cheung C, Chen H \(2010\) The Continuance of Online Social Networks: How to Keep People Using Facebook? In: 43rd Hawaii International Conference on System Sciences](#)
- [Singh N, Zhao H, Hu X \(2003\) Cultural adaptation on the Web: A study of American companies' domestic and Chinese websites. Global Information Management 11\(3\):63–80](#)
- [Snell W, Miller R, Belk S \(1988\) Development of the emotional self-disclosure scale. Sex Roles 18\(1\):59–73](#)
- Stengers H, De Troyer O, Baetens M, Boers F, Mushtaha A (2004) Localization of Web Sites: Is there still a need for it. In: International Workshop on Web Engineering (HyperText 2004 Conference)
- [Suler J \(2004\) The online disinhibition effect. CyberPsychology & Behavior 7\(3\):321–326](#)
- [Tidwell L, Walther J \(2002\) Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. Human Communication Research 28\(3\):317–348](#)
- [Tufekci Z \(2008\) Can you see me now? Audience and disclosure regulation in online social network sites. Bulletin of Science, Technology & Society 28\(1\):20–36](#)
- [Vasalou A, Joinson A, Courvoisier D \(2010\) Cultural differences, experience with social networks and the nature of true commitment in Facebook. Human-Computer Studies 68\(10\):719–728](#)
- [Vekiri I, Chronaki A \(2008\) Gender issues in technology use: Perceived social support, computer self-efficacy and value beliefs, and computer use beyond school. Computers & Education 51\(3\):1392–1404](#)
- [Walther J \(1996\) Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. Communication research 23\(1\):3–43](#)
- Wilson P, Nolla A, Gunawardena C, Lopez-Islas J, Ramirez-Angel N, Megchun-Alpirez R (2002) Effects of cultural differences on e-mail communication in multicultural environments. In: Proceedings of the 3rd international conference on cultural attitudes towards technology and communication, pp 51–68
- [Ying X, Wu X \(2010\) On link privacy in randomizing social networks. Knowledge and Information Systems pp 1–19](#)
- [Zhou B, Pei J \(2010\) The k-anonymity and l-diversity approaches for privacy preservation in social networks against neighborhood attacks. Knowledge and Information Systems pp 1–31](#)

## Author Biographies



**Slava Kisilevich** received the M.Sc degree in Information Systems from Ben-Gurion University of the Negev, Beer-Sheva, Israel, in 2008. He is currently working toward the PhD degree in the “Databases, Data Mining and Visualization” group of Prof. Dr. Daniel Keim, at the University of Konstanz, Germany. His research interests include Data Mining, Text Analysis, Information Retrieval, Recommender Systems, Geo-spatial Analysis, Visualization and Social Network Analysis.



**Chee Siang Ang** is a Lecturer in Multimedia and Digital Systems in the School of Engineering and Digital Arts, University of Kent at Canterbury. Before joining Kent, he was a research fellow at the Centre for Human Computer Interaction Design, City University London, where he completed his PhD in the area of social gaming. His main research interest lies in the general area of Internet study and social computing, specifically virtual worlds and online communities. He is particularly keen to understand and analyse human behaviour and social interaction in a virtual environment as well as social network structures that emerge from this interaction. A really fascinating aspect of such research is the emerging cyberspace such as MMOGs, 3D virtual worlds and online social networking and how we can exploit these tools in various contexts of applications, including training, healthcare, and behaviour intervention.



**Mark Last** is an Associate Professor at the Department of Information Systems Engineering, Ben-Gurion University of the Negev, Israel and the Head of the Software Engineering Program. Prior to that, he was a Visiting Research Scholar at the US National Institute for Applied Computational Intelligence, University of South Florida, USA. Mark obtained his Ph.D. degree from Tel Aviv University, Israel in 2000. He has published over 150 papers and chapters in scientific journals, books, and refereed conferences. He is a co-author of two monographs and a co-editor of eight edited volumes published by Springer-Verlag, World Scientific, and IOS Press. Prof. Last currently serves as an Associate Editor of IEEE Transactions on Systems, Man, and Cybernetics and Pattern Analysis and Applications (PAA). His main research interests are focused on spatio-temporal data mining, cross-lingual text mining, Social Network Analysis, and security informatics.

---

*Correspondence and offprint requests to:* Slava Kisilevich, Department of Computer and Information Science, University of Konstanz, Universitaets Strasse 10, Box 78, 78457 Konstanz, Germany. Email: slaks@dbvis.inf.uni-konstanz.de