

# "Ten Things I Want to Know About You": Unveiling Friends' Attribute Relevance

João Guerreiro, Daniel Gonçalves

Technical University of Lisbon / INESC-ID  
Rua Alves Redol, 9. 1000-029, Lisboa, Portugal.

joao.p.guerreiro@ist.utl.pt, daniel.goncalves@inesc-id.pt

## ABSTRACT

**Motivation** – The growing amount of personal information on the web raises increasing concerns about what and with whom we share information online. Nevertheless, little effort has been made in determining the relevance of the information shared with us or in filtering it accordingly. This is even more important considering our need to be constantly aware of what is happening in our friends' lives.

**Research approach** – A study to identify the most relevant characteristics when seeking information about friends and to scrutinize which specific features they mention. To achieve that, we resorted to interviews and questionnaires. We asked participants to describe people and asked them to rate the perceived relevance of a carefully pre-determined set of attributes.

**Findings/Design** – Results suggested that the most relevant attributes when seeking information about friends are: *Personality, Relationship, Interests & Hobbies, Academic History, Profession, Phone, Email and Address*. We also provide indications of the specific features people mention when referring these attributes.

**Take away message** – The relevance among attributes varies when seeking information about friends. It should be considered to warn users or highlight the changes when they occur in the most important attributes.

## Keywords

First-person relevant information, Social Relationships, User Study, Profile Attributes, Personal Information.

## INTRODUCTION

The web and the boost of Social Networks (SNs) have led to an increase in the amount and type of Personal Information (PI) available. Moreover, with the number of connections always rising the accessible information goes beyond the barrier of our closest relations (Lugano, 2008). This boost is followed by the frequent pursue for information about people. Search engines are regularly used for this purpose (Spink, 2004) and one of the most popular uses given to *Facebook* is social surveillance/investigation (Joinson, 2008) as users want to know more information about others or just keep *up-to-date*.

Friend-of-a-Friend (FOAF) is the most popular ontology used to describe people and their relationships

and is used as a resource for multiple applications (Powers, 2003). However, the existent ontologies do not seem to be based on a theoretical background defining which attributes to consider. The same happens with previous research focusing in obtaining information about people (e.g. Guerreiro, 2012; Zhou, 2012). In the overall, we do not know what kind of information people want to know about others.

SNs user profiles are a rich source of information as they comprise a large set of personal attributes. In fact, previous research acknowledged that these attributes differ in relevance whether to predict the number of friends in a SN (Lampe, 2007) or to form accurate impressions about people through condensed profiles (Stecher, 2008). Whereas these studies can provide insights of the most relevant attributes for their specific contexts, it is not possible to generalize to scenarios where we seek information about people.

There has been an increasing concern about privacy (Fogel, 2009) and the development of mechanisms to determine with whom we share our information. Understanding the way people organize their contacts (Kelley, 2011) improves the way they manage the *outflow* of information. However, little headway has been done to help them manage its *inflow*. Knowing exactly what information the users may want would decrease the effort to find relevant information. In fact, it would allow warning users (or highlighting) when some friend's important attribute changes (e.g. "John Smith is now working at INESC").

We performed a study where participants were asked to thoroughly describe people. In addition, we asked them to rate the relevance of a set of carefully selected attributes when searching information about people. The former provided insights about which characteristics people use when describing someone and the latter about the ones they find relevant when seeking person-related information. Their combination allowed us to get the characteristics' relevance for people with different intimacy levels. In this paper, we focus on the analysis of a specific relationship type, friends, due to our very frequent need to be aware of what they have been doing. Herein, we analyze the interviews and questionnaires about friends and show the most relevant attributes when we want to know information about them.

## “TEN THINGS I WANT TO KNOW ABOUT YOU”

We frequently seek information about someone and are usually presented with the same features, which relevance was not previously validated. Thus, we conducted a study that aimed to find out which are the most relevant attributes that we intend to see in these situations. The knowledge and intimacy levels differ among relationships, fact that we took into consideration when preparing the study. This paper aims at providing the **most relevant attributes** when **seeking information** about **friends**.

### Methodology

We interviewed 23 participants (14M/9F) with ages between 19 and 58 (M=32; SD=12.2). They were asked to describe different people, enabling us to find which features are usually mentioned when talking about someone. They also filled a questionnaire, where we collected additional information about the relevance of several attributes from the participants' perspective.

#### Describing People

In this phase we wanted to lay out the attributes people usually refer to when describing someone and find out eventual differences between distinct relationship types. After asking permission to record the interview for further analysis, we asked them to “*describe a close friend, providing both profile and previous interactions information*”. This was carefully worded to try to elicit as much information as possible. We performed five pilot interviews where we asked users to “*describe a close friend*” without supplementary instructions, which led to descriptions limited to physical and personality features. In contrast, by asking for profile information we were able to find the most relevant features for each person looking beyond basic facts. It includes life story and achievements that stood out, which may include personality details and personal experiences from different perspectives. Asking for previous interactions aimed at unveiling information about their relationship.

In addition to a close friend, participants described an acquaintance and a famous person. The order changed among users to prevent biasing the results. The main reason to describe a *Friend*, an *Acquaintance* and a *Famous person* was to identify the differences when talking about people we may have different knowledge levels (from intimate to generally-available information). This paper centres the results and contributions on *Friends*, due to the recurrent desire to be aware of changes regarding this relationship type.

#### Pre-Determined Characteristics

We selected a set of attributes which helped us to better catalogue the descriptions', easing the analysis and allowing reaching to information that participants knew but did not promptly mention (via specific questions). We based our selection in areas related to people descriptions, such as person ontologies (e.g. FOAF), SN profiles (e.g. *Facebook*) and previous studies on profile attributes (Lampe, 2007; Stecher, 2008). In addition, the 5 pilot interviews allowed us to collect other attributes.

The pre-determined features are: *Email; Address; Phone; Homepage; City; Birth Place; Workplace; Profession; Professional History; Academic History; Profession area; Interests & Hobbies; Age; Related People, Relationship user-person; Marital Status; Family Information; Highlights; Physical Description; Personality; Nickname; Peculiar Expressions; Religion; Political Ideology and Past Addresses.*

#### Collecting the Data – Spontaneous Versus Induced

We wanted to know the information that the participants found more relevant, so we tried to separate the information that they did not mention because they forgot from the one they did not actually know. Indeed, it might happen that users would get stuck when describing someone, but a gentle nudge might prompt them to continue, leading to more information. That is why we classified the information they revealed in two main categories: *Spontaneous*, when it is described by the participants without any intervention from the interviewer, and *Induced*, if it directly resulted from some question or remark on his part. The interviewer had a set of pre-determined questions that could be asked to try to elicit forgotten features. These questions were asked only when the participants had finished their spontaneous descriptions and were made in a way to extend them, instead of direct yes/no questions. Some examples are: “*What do you know about his past/origins?*” and “*Which activities or work do you share with him?*”. Besides *Induced* and *Spontaneous*, the attribute could also be classified as *Not known Spontaneous* (when participants refer the attribute but only to clarify they do not know it) or *Not Mentioned*.

#### Relevance Questionnaire

This questionnaire intended to know what people find relevant when needing information about others. We asked them to imagine they needed information about a close friend, an acquaintance and a famous person, and to classify the relevance of each of the pre-determined attributes, using a 5-point *Likert Scale*, ranging from 5-*Very Relevant* to 1-*Irrelevant*. We made clear that these ratings were about the general types instead of focusing on who they had described earlier. This step allowed us to identify attributes that users found relevant but did not mention on their descriptions either for inadequacy as profile information, not knowing it for that specific person or just forgot to mention it. We were also able to validate the relevance of attributes that were spontaneously mentioned in the interview.

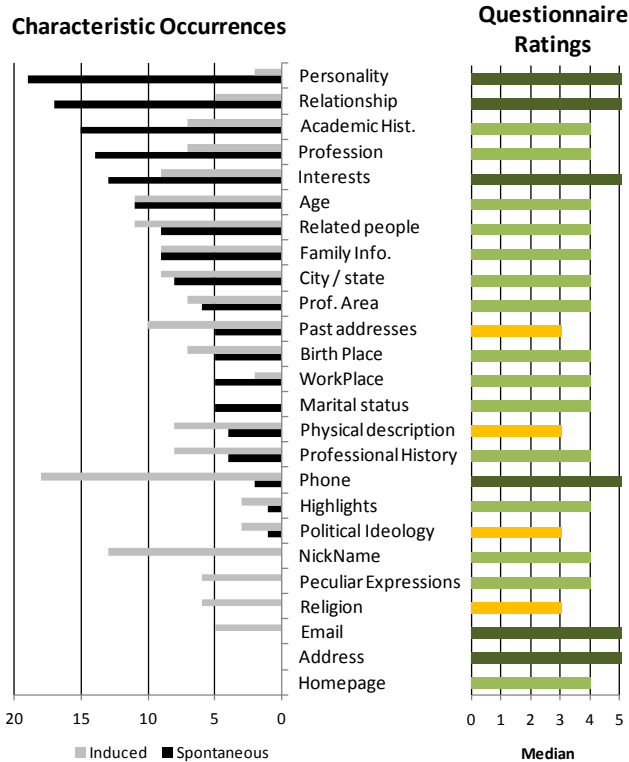
## RESULTS

Herein we present our study's results focusing on interviews and questionnaire ratings about *Friends*.

#### Friends' Attribute Relevance

Figure 1 shows the number of participants that mentioned each characteristic spontaneously or induced (left) and the median of questionnaire's ratings (right).

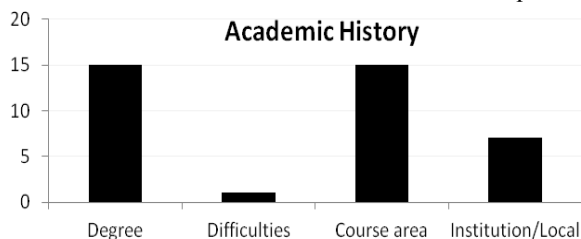
The former shows that, respecting *Friends*, the most popular attributes were mentioned by nearly the same number of participants but with differences on *Induced*



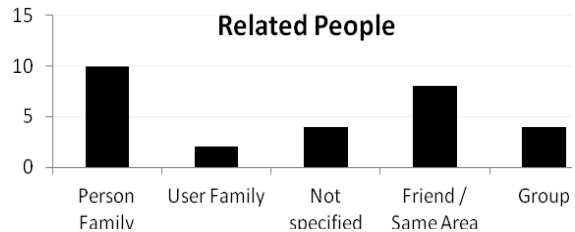
**Figure 1. Spontaneous and Induced frequencies for each characteristic (left) and the median classification of their relevance (right), in respect to Friends.**

and *Spontaneous* frequencies. It implies that they knew all these attributes about their friends, but did not mention some of them straightaway due to lack of relevance or forgetfulness. The leading attributes for *Friends* were: *Personality*, *Relationship*, *Academic History*, *Profession* and *Interests*, which were all mentioned spontaneously by more than half participants. *Age*, *Related People*, *Family Information* and *City* were also popular. In contrast, some attributes were never (or just a few times) mentioned spontaneously. No participant mentioned *Address* and *Homepage*, while *Email*, *Religion*, *Peculiar Expressions* and *Nickname* were only referred by induction. *Phone* and *Nickname* stood out among the less popular due to the references by induction. That means they knew that information but for some reason did not mention it.

The questionnaire shows that *Phone* and *Nickname* weak scores were not due to lack of relevance as they classified them as *Relevant* (*Nickname*) and *Very Relevant* (*Phone*). This suggests that users forgot to mention these features or did not find them adequate as



**Figure 2. Academic History detailed frequencies**



**Figure 3. Related People detailed frequencies**

profile information. Likewise, *Email* and *Address* were found *Very Relevant* (Median(MD)=5). The relevance of the most mentioned attributes were also supported by the questionnaire, mainly *Personality*, *Relationship* and *Interests* (MD=5). In fact, most features were classified *Relevant* (MD=4), but some less popular ones were classified as *Nor Relevant or Irrelevant* (MD=3), suggesting that their presence or absence is indifferent when seeking person-related information. These features are: *Religion*, *Political Ideology*, *Physical Description* and *Past Addresses*.

### Detailing Characteristics

We collected which specific features users mentioned for each characteristic. As an example, for *Interests & Hobbies*, participants mentioned mainly *sports*, *arts* and *technology*. That information was gathered from a typical yes/no counting of occurrences. In what follows, we detail the most relevant attributes with noteworthy results on the specific features mentioned.

### Academic History

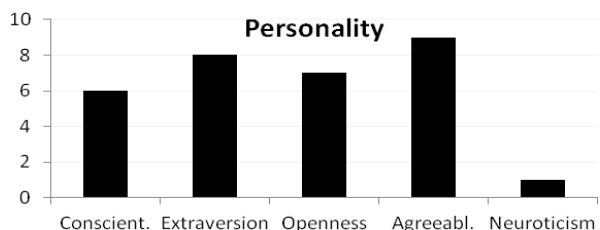
Regarding academic information (Figure 2), people often refer the degree and the area where they got (or are getting) the degree on (“He is taking a Master’s on Computer Sciences”). The institution and/or local are also common (“He studied at Instituto Superior Tecnico in Lisbon”). One participant referred the struggles that his friend encountered to finish the course/school (“After his father died, he encountered some problems finishing high school (...”).

### Related People

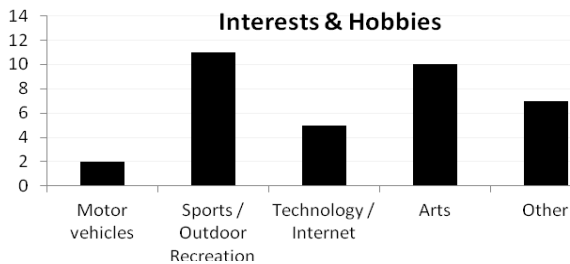
Users referred the person’s family often (Figure 3), which is explained by the proximity of a close friend. Mentioning shared friends or someone working in the same domain was also common (e.g “Works with me and John at (...”). Referring specific (or not) groups of people and the user’s own family was less frequent.

### Personality

We have based our classification for *Personality* on “Big 5 Factors” recognized by many researchers



**Figure 4. Personality detailed frequencies**



**Figure 5. Interests and Hobbies detailed frequencies**

(Digman, 1990): *Conscientiousness*; *Extraversion*; *Openness*; *Agreeableness* and *Neuroticism*. Figure 4 states that *Neuroticism* was the less popular, while the others are somehow balanced.

#### Interests & Hobbies

When talking about *Interests & Hobbies* (Figure 5), the majority resorted to *sports & outdoor activities* (e.g. hiking) and *arts* (eg. music, literature and cinema). *Other* mentioned interests were mainly comedy, astrology and travelling. Some examples of *Interest & Hobbies* descriptions are: “We used to play soccer together”; “He knows a lot about cars and planes”).

#### Relationship between User-Friend

When talking about the relationship with their friends, many participants (15) mentioned the type of relation they had (e.g. “He is my best friend”; “We are soccer teammates”). They also mentioned what they usually do or did together (12) and when/where they interact (11) (e.g. “I’m with him on Mondays on our soccer games”).

#### DISCUSSION

Resorting to both interviews and questionnaires, the most relevant features when seeking information about friends are: *Personality*, *Relationship*, *Interests & Hobbies*, *Academic History*, *Profession*, *Phone*, *Email* and *Address*. Most of these attributes are subject to change and, except *Personality*, are part of social network profiles. Their preponderance in this study suggests that it should be considered to provide mechanisms to warn users or highlight these changes when they occur.

Whilst most other features are also relevant and could benefit from similar (but weaker) measures, *Religion*, *Political Ideology*, *Physical Description* and *Past Addresses* were not referred many times and had lower ratings in the questionnaire, suggesting indifference.

#### LIMITATIONS AND FURTHER RESEARCH

We believe we provided valorous insights of attributes’ relevance in the person information domain, nevertheless we cannot argue a broad generalization. A higher number of participants would allow us to a better validation of results. We approached this subject with a focus on *Friends*. We believe these results can help distinguishing the most important features, in general, for this relationship type. We do not claim those are the best for all contexts. Nevertheless, we believe that in most contexts, the relevant features about *Friends* will be narrowed among the ones this study identified. It will

be one of our major and following concerns, as we will try to validate it and establish the relation between different contexts and relationship types.

#### CONCLUSIONS

The increasing data shared on the web, mainly on (but not restricted to) SNs, led to privacy concerns about what and with whom we share personal content. While this problem seems to be addressed little effort has been made to realize if all that information is relevant to us. We conducted a study that showed that not all attributes about friends are equally relevant. We have scrutinized the attributes’ relevance when seeking information about friends and which particular information they usually refer to. Applications or frameworks that make use of this kind of information, such as SNs can employ these findings to provide a better user experience, highlighting the most important features. It is not about what we share, but what we want to know about them. The fact that someone shared something does not mean we want to see it and more comprehensive results may be burdensome instead of helpful.

#### ACKNOWLEDGMENTS

Work supported by the Portuguese Foundation for Science and Technology (FCT): João Guerreiro’s scholarship, grant SFRH/ BD/66550/2009; project AAL/0014/2009; project PEst-OE/EEI/LA0021/2011.

#### REFERENCES

- Digman, J.M. Personality structure: emergence of the five-factor model. *Annual Review of Psychology* (1990)
- Fogel, J., Nehmad, E. (2009) Internet social network communities: Risk taking, trust, and privacy concerns, *Computers in Human Behavior* 25.
- Guerreiro, J, Gonçalves, D. (2012) Visualiz’em: “Show me more about him!”. *In proc. of AVI*.
- Joinson, A. N. (2008) Looking at’, “ Looking up ” or “ Keeping up with ” People ? Motives and Uses of Facebook. *Social Networks*, 5.
- Kelley, P.G. et al. (2011) An Investigation into Facebook Friend Grouping. *In Proc. of Interact*.
- Lampe, C., Ellison, N. and Steinfield, C. (2007) A familiar Face(book): profile elements as signals in an online social network. *SIGCHI*, ACM Press.
- Lugano, G. (2008) Mobile social networking in theory and practice. *First Monday* 13.
- Powers, S. (2003) Practical RDF. *O’Reilly & Associates*
- Spink, A., Jansen, B.J., Pedersen, J. (2004) Searching for people on web search engines. *Journal of Documentation* 60, 266–278.
- Stecher, K., & Counts, S. (2008) Thin Slices of Online Profile Attributes. *In proc. of AAAI*, 127-135.
- Zhou, M. et al. (2012) Finding someone in my social directory whom I do not fully remember or barely know. *In Proc. of IUI*. ACM, 203-206.