

Computerisation, data privacy and scientific excellence; Where are we going?

Damir Kalpić
Department of Applied Computing
Faculty of Electrical Engineering and Computing
Unska 3, 10000 Zagreb, Croatia
damir.kalpic@fer.hr

Abstract. *Computerisation of most human activities that handle data is widely in function nowadays. General fears of Big Brother syndrome are ubiquitous. The paper attempts to challenge this common wisdom by demonstrating on some examples how computerisation in connection with proclaimed data privacy politics actually hinders the useful information, some of which was even formerly available, without computerisation or in older information system versions before the current rigid data privacy legislation. Examples are taken primarily from some personal author's experience and from information systems aimed for academic community in Croatia, developed at the author's Department. Data privacy prevents ordinary members of the academic community or society from proper investigation that could challenge the adequacy of existing policies. At the same time, the persons who institutionally have granted access to the collected information and who are the most influential to conduct policies have hardly time to do so.*

Keywords. Personal identification number, Identity fraud, Risk, Student administration, Subsidised nourishment, Academic achievement.

1. Introduction

This invited paper offers the author a possibility to share with the audience some his first-hand experience in order to challenge some seemingly generally accepted threats, like disclosure of private data through computerisation. The famous personification of the ubiquitous presence of data collectors and abusers is named *Big Brother* after the Orwell's character in his well known futuristic novel *1984*. However, these widely expressed fears are in contradiction to the popularity of

the equally named TV reality show where some irrelevant individuals perform some irrelevant activities while being thoroughly watched by wide enthusiastic audience who find inexplicable joy in watching someone's alleged privacy. The press is overwhelmed with gossip and stories from private life of so called celebrities. This information is of no use to anyone, except that it can ruin someone's life. This duality of human character is appropriately exploited by so called data privacy activists. How much the transparency and tax payers' rights are endangered by hiding information which is of our common concern, is hardly ever asked.

On the other hand, overwhelming computerisation allows for providing to worldwide audience the measures of scientific activity for any scientifically active individual, whether s/he likes it or not. Under circumstances of scientific overproduction and overflow of publications, some measures have been introduced to help reduce the noise. However, these indices aimed for readers, have turned into major factors for evaluation of individual scientific excellence. This attitude shall be under scrutiny in this paper.

The topics are considered thanks to the author's experience regarding the abolition of the Personal Identification Number in Croatia, development at his Department of two nationwide information systems to support higher education and through his recurring duty of evaluation of his peers' proposals and performance for their re-election or promotion.

2. Personal Identification Number

In early 1970-ties the introduction of Unique Citizen's Identification Number started as an answer to requirements of early computerisation efforts. At that time, no computer networks were available, so that

rules regarding the construction of that number had to enable its creation locally. The number was determined to be 13 digits long. It started with the date of birth, where the year was presented by 3 digits, so it was not affected by the Y2K problem. Locality of birth code and sex code followed, ending with a serial number of the individual in the local birth registry. The last digit was a control one, calculated modulo 11. In the years to come, this number has become the commonly accepted unique key to identify the citizens. In 2002 the Croatian Parliament in a poor and misunderstood imitation of foreign practice, pretending to protect data privacy, brought a legal act practically abolishing the use of that allegedly “all-revealing” number. There were reasonable voices in the Parliament to object this Act but they were overridden by the computer-illiterate majority. An exchange of existing personal documents followed, having the number removed.

The legislators did not bother to consult profession. Afterwards, the author was engaged by the Government in a group of about 40 professionals from different fields to solve the problem how to proceed further without the Unique Personal Identification Number. After having achieved the e-mail addresses of all the Group members, the author sent them a longer e-mail message explaining that this number was used as a substitute primary key instead of the imperfect combination of family and personal names, extended by the date and place of birth and father’s name; as it had been usual beforehand. The author pleaded for withdrawal of the Act instead of trying to find solutions to an unnecessarily provoked problem. He also offered a modest financial reward to any information that anyone could obtain from his Personal Id, except his date of birth, place of birth and sex. After these mailings, the Group was never convened again. The author had an opportunity to criticise the Act at an open public discussion, in a radio talk and in form of an interview in the major daily newspaper [1]. In this interview he admitted that the Personal Id does unnecessarily reveal a personal information – date of birth, which was inevitable due to the technology available when it had been introduced. He proposed, if necessary, but hardly as a priority, a one time substitution with non-revealing numbers. The efforts of some other individuals and

organisations, like the Banking union also were in wane. The existing documents were substituted with useless ones. The citizens had to provide other documents as statehood certificate or birth certificate and/or to carry with them the old voided documents in order to prove their identity.

The Act has made it easier to steal identity or to hide own wrongdoing. Data privacy activists were happy to deprive the banks of a unique person’s identification. This made possible to a person deeply in debt and unable to repay it, to get another loan in another bank and double the damage. The banks’ normal reaction would be to increase the interest rates for loans as result of increased risk. Honest citizens can be punished due to villains. Eventually, the banks reintroduced the use of Personal Id. At first, it occurred “voluntarily”. A person, who would not submit it, would not be considered for a loan. Afterwards, the banks did not have to fear because the very Ministry of finance required the Personal Id for taxation purposes. Some other institutions were damaged. The author had the opportunity to witness it at a Croatian state agency which had problems to collect some fees from land owners who were not uniquely identifiable. There were identity frauds where a person was selling someone else’s real estates due to increased mess in cadastre and even more so in land registry.

To be fair, one must admit that even some honest people could take advantage of the Act, as illustrated in the following story:

The author’s colleague in rush to a business meeting had parked his car in the city centre at a prohibited place. After the meeting, his car was not there, obviously towed away by the communal service, just as he expected. He went to the tow lot to pay the fine and take his car back. The car was not there. There was no information about having towed away his car. The author’s colleague went back to the position he had parked the car in order to call the police and report his car stolen. But his car was back there! How come? The author’s colleague bears the same personal and family name as the newly appointed minister of transportation at that time. While checking the car owner without the Personal Identification Number, one confounded the author’s colleague with his boss, the new minister. The reaction was quite expected.

In the end of the day, this year, the so called Personal Taxation Number is to be introduced. It will not reveal any personal information but it will completely assume and extend the role of the former Personal Id. Nobody officially admits that it could have been done immediately, as the author proposed in [1]. Nobody in author's knowledge, at least until now, has come to the idea to complain about the endangered personal data privacy introduced through this new number. The author wonders if some data privacy activists could miss such a splendid opportunity to let themselves heard. Ignorant technophobes may again raise their voices against the "all revealing number", while if somebody bears a rare, uncommon name, which also uniquely identifies a person, nobody objects. The unhappy ones among honest citizens are actually those who have very common names and are often confounded with someone else.

3. Subsidized nourishment of students

Ministry of Science and Technology at the end of the last century faced an annoying problem. In major student feeding facilities or students' restaurants, food stamps were issued to students after a payment. The students were using these stamps to pay for their meals and the Ministry had subsidized this activity. It can be easily devised that the Ministry was paying also when nobody did actually eat and the subsidising money was split between student and the restaurant.

In order to avoid the misuse in this activity with the budget larger than the University's, the Ministry financed the project of a computerised system for management of students' nourishment based on a sort of credit cards issued to students [3]. The system has been successfully put into function and it is still running. One of the aims of the system was to detect inappropriate use of subsidise. Statistics showed some abnormality in the distribution of consumption. Some students did not use the system for what it was designed - to enable them regular convenient and appropriate food. Instead they were buying resalable packed products like juices, packages of e.g. 27 puddings or yoghurts etc. The developers of the system organised web sites where such abuses were made public. On the other hand, the system could reveal the cashiers who were selling such packages with

no questions asked. Very soon, revealing of this information was banned by the Ministry because "it would offend the abusers' rights". So, the abuses persist, probably lower than before the system introduction, but they still represent a misuse of the taxpayers' money and encourage socially unacceptable behaviour among the student population. Taxation money is used partially to support fraud. Honest students appear to be fools for not enjoying fully the "beauty" of the system.

4. Student Administration Software

The tradition of developing software to support student administration is rather long [4][5][6][7] but each version was becoming ever more restrictive following the legal regulations. The last version [8], aimed for the higher education in the whole country is strictly obeying the data privacy rules in order to satisfy the most restrictive user. How restrictive a user can be, illustrates the recorded case of legal suite of a professor against his colleague because he had examined the accuser's student in a course they commonly taught. Beforehand, some analyses, if even in their initial stage, had been preformed in order to try to find some behavioural patterns of students and educators [9]. Now, thanks to an unfortunate appeasement to pusillanimous behaviour of some faculty members already at the time of the former software version and due to fear from being accused for breaching the data privacy, excessive limitations have been introduced. The data privacy enforcement is omnipresent. The mentor of a student can only observe the grades his student has achieved in the mentor's courses and the overall average grade. Whether his or her student is more inclined towards mathematically oriented courses or hardware or software or social sciences, whether the student is a regular one or a notorious repeater, remains for the mentor a mystery. The mentor is also supposed to advise his mentored students regarding the enrolment in elective courses. Publicly visible, there exist the contents of each course. The professors are obliged to enter also the time plan for weeks of educational activities, but as they do not see it immediately, it is understandable that they are reluctant to do it, or they use alternative means to expose it. Even if the very precise educational schedule

and contents of an elective course were available, it is not enough. The course quality and attractiveness heavily depend upon the execution. There are regular anonymous students' evaluations of educators, three times per course, for each course separately [2]. The deviation from data secrecy is in the fact that all the educators in the same course can see each other's scores. The written students' comments are hidden, except to the educator in question, what is appropriate because the comments can sometimes be singular no justified attacks, while the scores are of statistical nature and represent fairly the audience's opinion. Still the information among all the teachers is not revealed.

This is a giant leap forward in comparison to recent times when the educators were "untouchable". The author remembers how the first attempts of these enquires were abolished as it turned out that a certain professor's lectures ranked as "hardly understandable". For the author and his team who provided software, that outcome had been the best proof that the software performs correctly! Now, this secrecy appears to be very humane towards the educators. It certainly is. However, what about the students? A mentor can advise them only according to some gossip to avoid certain educators. Is it correct, or it might be protracting prejudice?

The students who prolong their study excessively are required to pay tuition. Their parents are eager to know what they are paying for. Is their descendant in the first or in the last year of study, or somewhere in between? The same can be said for the taxpayers and for all the students because even if tuition paid, it is only a fraction of the real costs. The majority of the costs are still covered by the taxpayers' money. The spending per student differs heavily among faculties and their source is far from transparent.

At once, the results of examinations have become a secret, although there has been in the university regulations since ancient times that the examinations are public. The author's Faculty is an institution that boasts with its acknowledged quality, it is the only Croatian faculty to be internationally accredited and it is notoriously free of bribery, as different enquiries among the University students had demonstrated. One of its prides is the long tradition of entrance examination process consisting of performance in the secondary

school and achievements at the entrance examination [10]. Among other things, it is made public how many credits a candidate has brought from the secondary school and what were his or her achievements at the examination. Recently the Croatian Personal Data Protection Agency tried to convince the author's institution to hide all the data regarding the entrance examination and to reveal only the personal result to each candidate. Of course, this could open unlimited opportunity to fraud and arbitrary behaviour. Decades of building confidence and eliminating of any possibility for bribery or corruption were invested in order to acquire the status of a fair and trustworthy institution for students, since the very beginning, since their entrance examination. A civil servant tried to endanger it, wishing desperately to find some argument for the Agency's existence. In accordance to the Bologna process, the examinations are being performed simultaneously, in large groups and homework has been introduced. All this is applicable under presumption that the students would not be cheating, what is absolutely understandable at any renowned university in the world. Our students had to sign statement obliging them not to breach these regulations. However, recently a few students were caught to be cheating without any doubt. The Faculty decided to announce their misdeeds in public as a warning that an expulsion may follow. This act has caused incredibly fierce reaction from the media and from the data privacy protectionists. They regarded it as an intolerable abuse of data privacy. It has remained unclear what would they suggest instead – expulse the students silently, so that no one knows and can experience it only if caught in the same criminal activity; or should these students have obtained a special reward for cheating? The Student Administration Software collects interesting and useful data. Only a fraction of it is ever used or reviewed thanks to data privacy enforcement. The very few who are entitled to examine the data, can hardly find time, motivation or competence to do it. Therefore, the whole system, developed and deployed after substantial efforts and invested competence, is reduced to a few very ordinary administrative functions, performed just because they are a duty. If there is no feedback to the data one enters, there is not much enthusiasm left among the educators for

filling in, less than while filling in a taxation claim.

Who benefits from the data privacy in state-owned public universities? Before the Bologna process analyses regarding examiners' attitudes could have been interesting. Nowadays, the examiners have little to say. The grades are derived from the credits achieved in continuous, mostly written examinations and are interpreted after the ECTS rules, according to the Gaussian distribution. What remains to be seen are the individual students' grades. Who are the best, who fair, mediocre or poor? Wouldn't it be interesting for the industry to have access to these data in order to offer scholarships not only to the best and the brightest, but also to those who may fit best their needs? Data privacy protects these students from such offers. Actually, hardly anybody knows who the really good students are. The overall average mark does not reveal everything. Some who have very high marks in all courses are not apt for every job. It is also interesting to know whether the grades have been achieved in first trial or after multiple repetitions. Employers cannot offer adequate salaries to newly graduated ones taking into account also their academic performance, in comparison with the others. The employers or professionals from industry and business act sometimes as guest lecturers in some practically oriented courses in order to reveal their professional experience. In this way they risk to be accused for revealing some of their business secrets for the sake of spreading knowledge. However, if they ask their host at the Faculty for information about students they could recruit, they are turned down due to data secrecy. Who benefits? Benefit the students who protract their study endlessly, or pretend to be excellent while being mediocre. They are supported by taxpayers and by their parents, presenting them false or no data at all. Good students are punished while the worst are rewarded.

Finally, there is a very recent example of a received e-mail, the author could not answer without trespassing the ruling official secrecy. It shows what is regarded quite normal elsewhere and what we hinder through data privacy:

Dear Professor Damir Kalpic,

X Y has submitted an application to the Faculty of Graduate Studies at the University of British Columbia in the following area:

***Program: Master of Science in Computer Science
Application Start Date: September 2008***

X Y has indicated that you will be providing a reference report on his/her academic ability and qualifications. References may be submitted via paper or using the online eReference form.

...

For further assistance, please contact our office at [UBC Graduate Studies Online Application](#)

*Regards,
Faculty of Graduate Studies
University of British Columbia*

...

5. Scientific excellence

In very contradiction to the above mentioned proceedings in the Croatian higher education and public life, so called measures of scientific excellence are sometimes, but not consequently, widely accessible.

The Rugjer Bošović scientific institute in Zagreb has developed and maintains the Croatian Scientific Bibliography (CROSBI) database. It is a very useful application which should have, but it has not, avoided the need to re-declare the same data in different occasions, like promotion, application for scientific projects, applications for scientific novices etc. Everybody can observe someone's achievements and conclude about competences thereafter. Any change of data requires authentication and authorisation, exactly as it should be.

On the other hand, the Croatian Ministry of Science, Education and Sports maintains the Registry of Scientists in Croatia. The data privacy is very strongly enforced so that the only information one can obtain is the Personal Scientist's Id if the user submits the person's personal and family name and the date of birth. This information is generally useless except for administration purposes where in some reports and applications this number is required. It serves to nothing because not even the

scientific status of the person gets revealed. However, for some administrative purposes a certificate that one is in some scientific status can be required. According to regulations, as published on the Ministry's Web, such a certificate can be issued to a scientist after a written request. The author has recently submitted such a request, and as he received it a month and a half later, it was no more useful for him. The author often acted as referee for someone's election or promotion, for evaluation of scientific projects and programs, for evaluation of proposals of study, for evaluation of proposals for establishment of new institutions etc. The main factor in all these considerations is the people who should perform. No information about them was available from the Register, initially envisaged as the main and trustworthy source of information. Data privacy hindered practical use of it. The referee can rely only on the data that the proponents have submitted about themselves. Some guilt for the criticised Register may be with the author of this paper who was engaged in the computerisation project for this Register. He should have maybe influenced its more appropriate fate.

Worldwide, information about scientists is widely accessible in scientific indices. The most famous one, the Current Contents, was envisaged to help the readers find relevant stuff under conditions of overproduction of scientific articles. In the meantime, much of the world academic community, have become addicts of scientometrics. The scientometric indices were not intended for the purpose of personal evaluations and the proof of someone's identity does not exist. The author can find in famous *Ovid* under "Kalpic, D" only his references. There is another Kalpic, "Kalpic B", unrelated and personally not known to the author, but who publishes in a close field as the author of this paper. If his name had started also with "D", the author (and his counterpart) would probably enjoy a higher reputation in some scientific circles! The information which is public is misused for the purpose it was not foreseen for. A wider discussion about the criteria for academic promotion deserves another paper. In brief, the author strongly believes that exclusive insistence on scientometric quantifiers is contra productive, driving some of the best & brightest towards irrelevant, but easy to publish niches of science.

6. Possible further developments

If the data secrecy is rigidly enforced further on, there will be someone to fill the gap, mostly with disinformation, of course. Nowadays, junk mail arrives in hundreds messages per day to those of us who have our e-mail addresses openly exposed. These are well known non customised messages related to easy money, sex, medications or replica watches. The sources of these fraudulent messages seem to be uncontrollable. It is conceivable that someone can on order "customise" such messages and send abusive contents regarding some person to many addresses. This could become a blackmailing instrument under circumstances that there were no publicly available sources of correct personal information. Recently, at the author's institution some alleged student posted on the Web the story about a professor being permanently drunk. The professor in question is an extremely sober and morally straight person but he was still injured by such accusations. How come that he was so wrongly accused? His family name can be reduced to a nickname meaning "cork" in Croatian. Someone who had only heard that nickname invented and posted the rest.

Junk mail and viruses have deteriorated our life on-line. The same can be expected to happen with disinformation. When it comes to absurdity, the legislators will have to admit their error and revoke the unnecessary information ban. It would be better if this can occur before too much damage has been done.

7. Conclusion

Data privacy policies should be seriously reconsidered, not only in Croatia. Outcomes of public education, paid by taxpayers' money, should be equally transparent to citizens like the results of scientific production which is often financed in the same way.

Internationally recognised scientific indices, widely open to public, should be used primarily as a valuable tool for information retrieval rather than measure of someone's scientific excellence.

Lack of reliable information can lead to disinformation and harm principally the correct citizens.

The author might have been too provocative, self-centred and abusive for an invited speaker. Nevertheless, the author strongly believes that at least some of his hypotheses do have foundation and he would be more than happy if this article could make any contribution to future improvements, or if a discussion could convince him that he was wrong. If proven to be wrong, there will be still time enough to prevent broadcasting the Peking 2008 Olympic Games results in public. The competitors' data privacy rights would be best protected if they were the only ones to know their own results.

8. References

- [1] Kalpić, D., Širanović, V.: I vladina grupa za matični broj (*Even a Government group for Personal Identification Number*), Večernji list, Zagreb, 22.01.2003. (interview).
- [2] Hoić-Božić, Nataša; Mornar, Vedran. Authoring of Adaptive Hypermedia Courseware Using AHyCo System // *Advances in Web-Based Education: Personalized Learning Environments* / Magoulas, George ; Chen, Sherry (ur.). Hershey, PA : IDEA Group Publishing, 2005. pp. 253-276.
- [3] Mornar, V., Fertalj, K., Kalpić, D., Krajcar S: Credit Card System for Subsidized Nourishment of University Students (2002), In Khosrow-Pour, M., ed.: *Annals of Cases on Information Technology 2002/4*, Idea Group Publishing, Hershey, PA, USA, ISBN 1-930708-40-8, pp. 468-486
- [4] D. Kalpić, J. Anzil, V. Dumanić, H. Zoković: Student Administration System, Proceedings of the 15th International Conference on Information Technology Interfaces, Pula, June 15-18, 1993, pp 123-128
- [5] D. Kalpić, H. Zoković, J. Anzil, D. Pukljak: Graphical Interface for the Student Administration System, Proceedings of the 16th International Conference on Information Technology Interfaces, Pula, June 14-17, 1994, pp 325-330
- [6] D. Kalpić, V. Mornar: Student Administration System, European review conference proceedings "University-Enterprise Information Systems" Graz, September 15-16, 1994, pp 124-131
- [7] Žepić, D. Kalpić: Student Service Information System, Univac Users Association/ Europe, Spring conference, München, 1983.
- [8] Kalpić, Damir ; Baranović, Mirta ; Mornar, Vedran ; Krajcar, Slavko. Development of an Integral University Management System // *International Conference on System Engineering, Communications and Information Technologies, ICSECIT 2001 Proceedings* / Magaš, Bedrich (ur.). Punta Arenas, Chile : Universidad de Magallanes, 2001.
- [9] Javor, Davor ; Statistička analiza baze podataka studentske službe (*Statistical analysis of the students administration database*), M.Sc. Thesis, Faculty of Electrical Engineering and Computing, Zagreb, 1996.
- [10] A. Žepić, P. Maćašović, M. Đurek, D. Kalpić: Primjena elektroničkog računala na klasifikacijskom ispitu (*Computer application for classification exam*), Informatica, Bled, 1974.