Teenagers need to aware their personal responsibility in internet society: our approach in information entrance course*

Yuanxin Ouyang*, Shanbo Cui*, Yue Qu*, Shiyuan Xiao*, Wenge Rong*, and Zhang Xiong*

*School of Computer Science and Engineering,Beihang University P.R. China,Beijing, China
Email:{oyyx,cuishanbo,quyuescse,xiaoshiyuan,w.rong,xiongz}@buaa.edu.cn

***Abstract*—The information literacy that the teenagers should have is widely concerned nowadays. We have been teaching a course entitled Introduction to Computer Science and computer Ethics that is taken by freshmen from computer science, electronic engineering and software engineering. There were more than 400 students in our course annually. The students were divided into groups to participate the in-class discussion. We used to focus on computer professional code of practice in last years. While in the fall semester 2017, we paid more attention on what computer and internet have bring us, and what should we do when we are facing the situations nobody had met before. There were 4 sessions, respectively great changes in society made by computer technology and internet, influences in our traditional culture, responsibility of enterprise in information age, and the artificial intelligent ethics problems. Most topics include new technology or idea. We got over 300 questionnaires after the course finished. From which we found that, compare with the former students, more moral topics have been closely followed. Nowadays, all society members are the contributors of the internet, whatever their major are. Our approach can be spread to similar information entrance course easily.**

***Index Terms*—information literacy, computer ethics, information entrance course, in-class discussion**

I. INTRODUCTION

Computer Technology and Internet have made great changes in our society, particularly in developing countries like China. Even people who lived in the areas that had never been reached by wired telephone service, now can access the Internet. In the past two decades, most computer science educators focused on the computer science researchers and software engineers' responsibility as well as their professional code of practice. But in recently years, more and more people aware that the safety and harmony of our information society depends on every members' self-discipline.

By the end of 2017, the China's national standard of information technology course in high school had been revised. According to the new edition, every high school graduate should have four basic information literacy, respectively sense and judgement of information, computational thinking, learning and innovation in digital environment and conscious-ness of their responsibility as an information society member, which includes [1]:

- Awareness and control of information safeness.
- Comply with information laws and regulations.
- Comply with code of ethics in information society.
- Comply with public standards in both real and internet society.
- Protect lawful personal profit, other peoples' profit, and public information safeness in relevant activities effectively.
- Care about environment problems and humanity problems brought by information technology revolution.
- Have the ability of initiative study, reason judgement, and responsible behavior, when facing new opinion and new matter created by information technology.

II. OUR APPROACH

Our course entitled Introduction to Computer Science and computer Ethics is taken by freshmen from computer science, electronic engineering and software engineering. About 400 students enrolled in our course annually. The students were divided into 4 or 5 member groups to participate the in-class discussion or debate. Each year, we usually design approximately 20 topics for the student to draw by lot.

We used to focus on computer professional code of practice in last years. Since more and more people have aware that computer ethics, as well as internet ethics should be understood and obeyed by every member of society, in the fall semester 2017, we paid more attention on what computer and internet have bring us, both positive and negative, and what should we do when we are facing the situations nobody have met before.

There were 4 sections in the last semester, respectively great changes in society made by computer technology and internet, influences in our traditional culture, responsibility of enterprise in information age, and the artificial intelligent ethics problems. In each section we offered 4 to 8 topics. Most topics include new technology or idea, such as driverless car, blockchain, bill-less payment, and personal recommendation system, their positive effect and the related moral or social

This work was partially supported by the National Natural Science Foundation of China (No. L1724045), and the Education Research Projects of Beihang University.

978-1-5386-1174-6/18/$31.00 ©2018 IEEE

problems. To meet the characteristic of 16 to 18 year freshman, some topics are a little bit "crazy".

A. Section 1 great changes in society made by computer technology and internet

- Advantage and disadvantage of bill-less payment
- Advantage and disadvantage of personal recommend-ation systems
- Nowadays, we can use our intelligent mobile phones anytime and anywhere, which makes the information we got become fragment, as well as our time management. Have we really improved our work or study efficiency?
- Internet and social networks make our privacy disappear gradually. Does the privacy disappearance beneficial or not beneficial to the development of our society?
- One university of offered a senior-level course called "Computer Viruses and Malware.", in which the enrollers will be taught how to write viruses, worms, and Trojan horses, as well as the history of computer viruses and how to block attacks. All the experiments were done within a closed computer network isolated from the Internet. While some computer security experts and researchers criticized that "No one argues criminology students should commit a murder to understand how a murderer thinks" [2] [3]. Debate whether the university was wrong to offer this course.
- Have Assange and Snowden obeyed the computer ethics?

B. Section 2 influences in China's traditional culture

- In traditional Chinese thinking, the sons and daughters should not hide any idea or opinion to their parents. Will you block your parents when sharing your moments in social networks like Weibo or Facebook?
- With the development of new information technologies like internet, virtual reality and augmented reality, the relationship between people will become more closed or sperate?
- What can Sharing Economy bring us in the so called "internet plus every industry" age?
- What is the boundary between public supervision and internet violence?

C. Section 3 responsibility of enterprise in information age

- Have you ever read the long long license terms when installing a software or accepting an internet service? Do you agree that these terms should be supervised to avoid unfair or illogical to the customers?
- The internet operation companies and service companies get all kinds of privacy data (e.g. location and photo) of almost every internet users. Do they have the right to use these data for other business logically (e.g. to push personal recommendation)?
- Is it morally acceptable to shut down a Web server that distributes child pornography by denial-of-service attacking? [2]
- Discuss about the technical methods that we can choose to keep the social society healthy and harmonious, analyze their usage and limitation.

D. Section 4 artificial intelligence ethics problems

- Ethics problems in using chitchat system (e.g. Microsoft Xiaoice).
- Ethics problems in using unmanned automobiles.
- The advantage and disadvantage when teachers and doctors were replaced by artificial intelligence systems.
- Will the popularize of artificial intelligent enlarge the gap between rich and poor.
- Is emotional intelligence needed to be developed? If the future artificial intelligence do have the same emotional intelligence with human being, would you like to establish close personal relationship with it?
- Should robots with artificial intelligence be given human rights? If yes, should they sacrifice when protecting human being?
- What kind of ethics problems will appear, when somebody's human memory and thinking ability were copied to an intelligent robot?
- Will the artificial intelligence destroy the human being?

III. QUESTIONNAIR ANALYSIS

A. Data Collection

This project used questionnaire survey method to collect data, and questionnaire survey method is currently the mainstream data collection method for user experience data. The questionnaire is designed based on the hypothesis model, and questions are asked from the perspective of user experience perception for the hypothesis of each hypothesis factor and the assumption of the role relationship.

The questionnaire was continuously improved through exploratory factor analysis. The final version of the questionnaire consisted of 22 topics and was divided into 4 attribute evaluations and one general evaluation. It mainly focused on students' evaluation to problems related to computer science ethics, and attempted to explore and analyze students' Which type of computer science ethical issue is of greatest interest. The questionnaire uses a 7-point scoring system, 1 is divided into "strongly disagree" and 7 is "resolutely endorsed". The degree of endorsement increases with the increase in scores. For 22 topics, the scoring data is also the data basis for the subsequent research. Our qualitative analysis questionnaire collected a total of 390 data, which requiring data clean-up. The basis for data cleansing is: The results of the questionnaire with single linear answers or regular answers are nullified. We cleaned up 72 poor quality data. After the data was cleaned up, a total of 318 sets of valid data were used to compose the initial data set for qualitative research. In the field of human-computer interaction for qualitative analysis, if the size of the data set is 10 times the number of items, then it can be viewed as an effective dataset [4].

B. Exploratory Factor Analysis

It is necessary to carry out exploratory factor analysis. The project uses exploratory factor analysis methods to determine the basic structure of the factors that affect the students' evaluation topics, and to solve the problem of strong correlation between variables in multivariate statistical analysis. Exploratory factor analysis is mainly used to determine the factors affecting the observed variables, the degree of correlation between each factor and each observed variable, and factor loading is used to intuitively infer the factor structure of the data [5]. We use the principal component analysis method to extract factors whose eigenvalues greater than 1, and classifying to four categories with 22 topics. The classification is shown below. The item No. refers to the Section No. and the topic No. introduced above.

TABLE I
RESULTS OF FACTOR ANALYSIS: MORAL ATTRIBUTES

Dimen -sions	Items	Mean	Std. Deviation	Load -ing	Cronbach's Alpha	KMO
		5.21	1.51			0.912
V101		5.16			0.848	
	D-1	4.82	1.55	0.459		
	D-2	4.69	1.54	0.522		
	D-3	4.97	1.44	0.485		
	D-4	4.89	1.51	0.465		
	D-5	5.59	1.61	0.602		
	D-6	5.60	1.52	0.533		
	D-7	5.66	1.47	0.669		
	D-8	5.06	1.66	0.540		
V102		5.36			0.867	
	C-3	5.54	1.525	0.489		
	B-4	5.67	1.374	0.470		
	A-4	5.31	1.371	0.552		
	C-1	4.98	1.549	0.665		
	C-2	5.13	1.641	0.726		
	A-5	5.40	1.708	0.659		
	A-6	5.48	1.564	0.577		
V103		4.49			0.811	
	C-4	4.65	1.495	0.549		
	A-1	4.25	1.515	0.699		
	A-3	4.66	1.602	0.591		
	A-2	4.52	1.460	0.616		
	B-3	4.36	1.523	0.567		
V104		5.85			0.671	
	B-2	5.78	1.286	0.647		
	B-1	5.92	1.351	0.661		

Usually we think that if the load is lower than 0.4, it means that the factor is not enough to represent a part of this variable and should be discarded. The load of all factors in the data obtained in this experiment is greater than 0.4, so the experimental data can be considered as a good data.

As for cronbach alpha, according to the definition of the critical value of this indicator by Murphy and David Shofer in 1988 [5], the criteria for the determination are shown as table V:

Sufficiency test is to test the accuracy of the data. Kaiser-Meyer-Olkin (KMO) method and Bartlett's spherical test are mainly used to verify the adequacy of the data. The KMO value can be used to test the data adequacy of multivariate statistics. The KMO value mainly reflects the comparison between the sum of squares of simple correlation coefficients

TABLE II
RESULTS OF FACTOR ANALYSIS: SOCIETY ATTRIBUTES

Dimen -sions	Items	Mean	Std. Deviation	Load -ing	Cronbach's Alpha	KMO
		5.45	1.42			0.918
V201		5.42			0.827	
	D-3	5.43	1.392	0.515		
	D-4	5.54	1.451	0.565		
	D-5	5.54	1.478	0.643		
	D-6	5.40	1.439	0.541		
	D-7	5.45	1.441	0.561		
	D-8	5.19	1.661	0.512		
V202		5.16			0.799	
	D-2	5.08	1.496	0.485		
	C-4	5.38	1.267	0.608		
	A-1	5.10	1.506	0.653		
	A-3	5.13	1.442	0.632		
	A-2	5.04	1.40	0.594		
	B-3	5.22	1.494	0.565		
V203		5.51			0.714	
	D-1	5.19	1.389	0.459		
	C-3	5.68	1.491	0.383		
	C-2	5.55	1.435	0.57		
	A-5	5.55	1.411	0.595		
	A-6	5.61	1.456	0.602		
V204		5.7			0.752	
	B-4	5.86	1.26	0.53		
	B-2	5.71	1.232	0.572		
	B-1	5.85	1.287	0.537		
	A-4	5.63	1.273	0.447		
	C-1	5.44	1.459	0.463		

TABLE III
RESULTS OF FACTOR ANALYSIS: TECHNOLOGY ATTRIBUTES

Dimen -sions	Items	Mean	Std. Deviation	Load -ing	Cronbach's Alpha	KMO
		5.40	1.37			0.858
V301		4.66			0.807	
	B-4	4.69	1.499	0.635		
	A-3	4.82	1.48	0.636		
	B-2	4.94	1.457	0.601		
	B-1	4.11	1.594	0.67		
	C-1	4.75	1.503	0.634		
V302		5.55			0.728	
	D-3	5.54	1.328	0.549		
	D-5	5.82	1.291	0.538		
	D-6	5.30	1.396	0.572		
	D-7	5.64	1.429	0.656		
	D-8	5.48	1.357	0.569		
V303		5.42			0.691	
	C-3	5.12	1.384	0.585		
	C-4	5.70	1.233	0.519		
	A-1	5.40	1.313	0.605		
	A-2	5.44	1.296	0.596		
V304		5.37			0.671	
	D-4	5.17	1.367	0.516		
	A-4	5.14	1.352	0.527		
	B-3	5.74	1.215	0.583		
	C-2	5.42	1.319	0.668		
V305		5.44			0.587	
	A-5	5.74	1.276	0.538		
	A-6	5.15	1.572	0.651		
V306		5.33			0.498	
	D-1	4.90	1.365	0.761		
	D-2	5.78	1.177	0.493		

TABLE IV
RESULTS OF FACTOR ANALYSIS: IMAGINATION ARRTIBUTES

Dimen -sions	Items	Mean	Std. Deviation	Load -ing	Cronbach's Alpha	KMO
		5.38	1.42			0.910
V401		4.97			0.876	
	C-3	4.93	1.495	0.515		
	C-4	4.66	1.453	0.508		
	A-1	4.86	1.472	0.534		
	B-4	4.91	1.492	0.589		
	A-3	4.94	1.552	0.694		
	A-2	5.10	1.345	0.44		
	B-2	5.09	1.406	0.603		
	B-1	4.72	1.576	0.442		
	A-4	5.28	1.368	0.468		
	B-3	5.15	1.494	0.452		
V402		6.00			0.72	
	D-5	6.16	1.165	0.586		
	D-6	5.90	1.308	0.491		
	D-7	6.11	1.268	0.629		
	D-8	5.90	1.427	0.501		
V403		5.22			0.693	
	D-1	4.89	1.433	0.400		
	D-2	5.16	1.347	0.667		
	D-3	5.34	1.409	0.663		
	D-4	5.49	1.422	0.506		
V404		5.34			0.715	
	C-1	4.91	1.484	0.522		
	C-2	5.20	1.472	0.617		
	A-5	5.71	1.378	0.498		
	A-6	5.54	1.395	0.55		

TABLE V
CRITERIA FOR THE DETERMINATION

Cronbach alpha Interval	Reliability
>0.9	High level
[0.8,0.9]	Moderate to high level
[0.6,0.8]	Low level
<0.6	Unacceptable level

and the sum of squares of partial correlation coefficients among various data variables [6]. The KMO value range is (0,1). The closer the KMO is to 1, the more significant the simple correlation result is compared to the partial correlation, and the data variable is more suitable for factor analysis calculation. However, close to 0 means that the correlation is too weak, to do factor analysis. In the KMO metric [6], a value of [0.9,1) is very suitable; [0.8,0.9] is suitable; [0.7,0.8] is general; [0.6,0.7] is not suitable; 0.6 or less is extremely inappropriate. Bartlett Testing is a criterion for calculating whether the correlation between one variable and another is large enough to allow factor analysis. If this variable is not relevant to other variables, then it cannot be used. With these variables, the Bartlett sphere test can be performed on the premise that the correlation between the correlation matrix and the identity matrix is significant and observable [6].

The experimental adequacy calculation results are shown in the table I to table IV, all of the data meets the requirements, indicating that the data is suitable for factor analysis.

As can be seen from the above tables, among the four attributes, the relevant score of society attributes is much higher than the other three attributes, indicating that the students of this course are most concerned about the social issues in the corresponding topics. The variance of technical and technological attributes is the smallest, indicating that the students of this course generally have relatively consistent views on the technical and technological levels of these issues.

IV. FEEDBACK TOPIC EXTRACTION

We got approximately 1000 feedback messages after taking back our questionnaires. These messages are mainly about research hot spots, social hot issues and the feelings of the class. To find out what our students mainly interested in, we used LDA (Latent Dirichlet Allocation) [7] to extract topics and keywords of these messages.

As a classic topic model, LDA is often used for inferring text topic distribution, which can be really helpful when arranging files and making summaries. Moreover, the topics generated by the LDA can be new features that can be used in other natural language processing missions.

LDA is a mature algorithm that there are many existed tools which have already achieve it. And gensim [8] [9] is a representative of this kind of tools. Thus, we got 20 topics with 200 keywords from these messages using LDA model (as shown in Table I). After deleting some meaningless words such as 'we', we make statistics on these words. Finally, according to the frequency of these words, we draw some Tag Clouds to make our results visible (as shown in Fig.1).

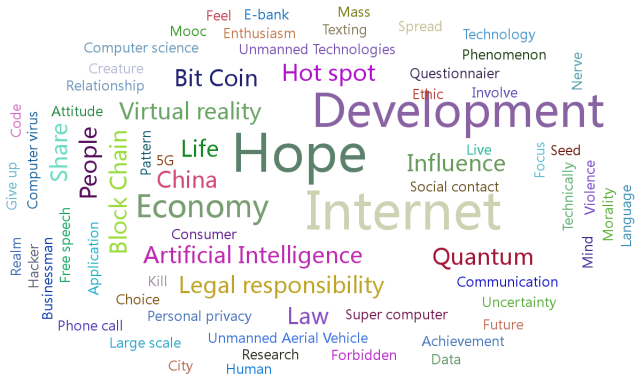

Fig. 1. Tag Cloud of our feedback keywords.

V. CONCLUSION

After questionnair data analyzing, we can classify the topics into different categories according to different attributes, Furthermore, we can judge students' interest in different attributes of different topics. And we extract principal components with those whose eigenvalue greater than 1 and correspond to the corresponding factors. Each factor corresponds to at least two topics, which ensures the validity of the factor. At the same time, several indicators are used to prove that the classification of this experiment is relatively reasonable. From which we can find out what kinds of topics and what kinds of keywords are included in the students' interest, and help them aware the individual responsibilities that teenagers, especially college students, should undertake in the Internet society.

REFERENCES

[1] Ministry of Education of the People's Republic of China. China's national standard of information technology course in high school. December, 2017.

[2] Quinn M J. Ethics for the Information Age: Global Edition[J]. 2015.

[3] Read B. How to write a computer virus, for college credit (cover story)[J]. Chronicle of Higher Education, 2004, 50(19): A33.

[4] Legris P, Ingham J, Collerette P. Why do people use information technology? A critical review of the technology acceptance model[J]. Information management, 2003, 40(3): 191-204.

[5] Peterson R A. A meta-analysis of Cronbach's coefficient alpha[J]. Journal of consumer research, 1994, 21(2): 381-391.

[6] Martin F, Ertzberger J. Here and now mobile learning: An experimental study on the use of mobile technology[J]. Computers Education, 2013, 68: 76-85.

[7] Blei D M, Ng A Y, Jordan M I. Latent dirichlet allocation[J]. Journal of machine Learning research, 2003, 3(Jan): 993-1022.

[8] Rehurek R, Sojka P. Software framework for topic modelling with large corpora[C]//In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks. 2010.

[9] Řehůřek R. Scalability of Semantic Analysis in Natural Language Processing[D]. Masarykova univerzita, Fakulta informatiky, 2011.