



**HAL**  
open science

## Definition of requirements for accessing multilingual information opinions

J. Derkacz, Mikolaj Leszczuk, Michal Grega, Arian Kozbial, Fernando Jorge Hernandez, Amaia Mendez Zorrilla, Garcia Zapirain Begona, Kamel Smaili

► **To cite this version:**

J. Derkacz, Mikolaj Leszczuk, Michal Grega, Arian Kozbial, Fernando Jorge Hernandez, et al.. Definition of requirements for accessing multilingual information opinions. *Multimedia Tools and Applications*, 2017, 10.1007/s11042-017-4737-3 . hal-01531313

**HAL Id: hal-01531313**

**<https://hal.science/hal-01531313v1>**

Submitted on 1 Jun 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Definition of Requirements for Accessing Multilingual Information Opinions

Jan Derkacz · Mikołaj Leszczuk · Michał Grega · Arian Koźbial · Fernando Jorge Hernández · Amaia Mendez Zorrilla · Begoña García Zapirain · Kamel Smaili

Received: July 31, 2016 / Accepted: date

---

Research work co-funded by the National Science Center, Poland, conferred on the basis of the decision number DEC-2015/16/Z/ST7/00559.

J. Derkacz  
AGH University of Science and Technology  
Al. Mickiewicza 30, PL-30059 Kraków  
E-mail: derkacz@kt.agh.edu.pl

M. Leszczuk  
AGH University of Science and Technology  
Al. Mickiewicza 30, PL-30059 Kraków  
E-mail: leszczuk@agh.edu.pl

M. Grega  
AGH University of Science and Technology  
Al. Mickiewicza 30, PL-30059 Kraków  
E-mail: grega@kt.agh.edu.pl

A. Koźbial  
AGH University of Science and Technology  
Al. Mickiewicza 30, PL-30059 Kraków  
E-mail: ariankozbial@gmail.com

F. J. Hernández  
University of Deusto  
Avda de las Universidades 24, ES-48007 Bilbao  
E-mail: fernandojorge@deusto.es

A. Mendez Zorrilla  
University of Deusto  
Avda de las Universidades 24, ES-48007 Bilbao  
E-mail: amaia.mendez@deusto.es

B. García Zapirain  
University of Deusto  
Avda de las Universidades 24, ES-48007 Bilbao  
E-mail: mbgarciazapi@deusto.es

K. Smaili  
University of Lorraine  
Campus scientifique BP 239, FR-54506 Vandoeuvre-lès-Nancy Cedex  
E-mail: Kamel.Smaili@loria.fr

**Abstract** In this paper, we will present a study concerning the understanding of the needs of people using Internet in order to access to multilingual information. In fact, in the framework of AMIS (Accessing Multilingual Information and opinionS), a Chist-Era project, we propose to develop a system which will help to understand the main idea of a video in a foreign language. In order to design a useful system, a survey allowing to specify the profile of potential users of AMIS has been conducted. The study concerned 170 people from different countries: Poland, Spain and France. The sample is composed of people of different ages and different culture and languages. The results, in terms of requirements, achieved from this study show differences depending on how often the people watch the news on TV or review them on the Internet, and on the age of the target group. These concrete results help us in several decisions concerning how to build a realistic architecture of AMIS.

**Keywords** opinion mining · social networks · multimedia analysis · multimedia accessibility · multilingual information · mother tongue language · end user · foreign language

## 1 Introduction

Nowadays, the information is widely available on the Internet under different forms and different languages. People communicate by posting information on social networks, by recording and sharing videos, by writing long texts, by recording audio messages and so on. Furthermore, this information is produced in several languages. Consequently, a new issue arises: how to make people accessing to a maximum of information pass the barrier of the language? Indeed, even highly educated people rarely speak sufficiently more than two or three foreign languages, while most know only one, and this significantly limits the access to news, culture, ... Permitting people to access to different information sources, would help them to have their own opinion about a topic by cross-checking the information. This challenge constitutes the objective of *AMIS: Access Multilingual Information opinionS*, a Chist-Era project, which, among other, has the ambition to present an automatic summary of a video to the end user in his mother tongue.

In order to understand the needs of the users, in terms of summarized information extracted from videos, we decided to study the requirement analysis of the end users. In fact, the access to information on the Internet depends on several parameters: age, gender, tongue mother, educational level, etc. This investigation helped us to understand what is expectable in accordance to the different attributes presented above and how the summary should be constructed. Building automatically a summary of a video presented in the mother tongue of the end user necessitates several components: video summary, Text summarization, audio summarization, extraction of overlaid text, automatic speech recognition and machine translation. The complexity of the architecture of such a system makes the requirement analysis necessary; one should really not be mistaken in the specification to succeed the realization of the automatic summarization of videos.

To do so, a large panel of people (170) from different countries: Poland, Spain and France have been surveyed. These people are from various cultures and different ages. In this paper, a detailed statistical study based on this panel is proposed to understand what kind of information has to be summarized into the mother tongue of

the end user. The remainder of this paper is structured as follows. After Related works Section, user requirements will be defined in Section 3. Based on this functional and architectural requirements, recommendations will be provided in Section 4. Finally, Section 5 summarizes an impact of expected results.

## 2 Related Works and Expected Progress Beyond State of the Art

The analysis of patterns of Internet and traditional news media usage were performed already in the first decade of widespread access to Internet.

One of the key questions was: to what extent can web become a supplement to television (and newspaper) news, or a substitute for these media? Based on a survey [8] of undergraduate students at a large public university where Internet was used on regular basis, this study implied that use of the Web as a news source was positively related with reading newspapers but has no relationship with viewing television news. Members of the surveyed community used the Web mainly as a source of entertainment. Patterns of Web and traditional media exposure were examined considering, i.e. desire for control and knowledge of political and societal issues.

This study indicates that even when computer skills and Internet access become more widespread in the general population, use of the Internet as a news source seems unlikely to diminish considerably the use of traditional news media. News consumption patterns were analysed within a specific group of people: members of university community, 520 undergraduate students, which is not representative for the larger adult population, and the patterns of media use in this university's student population can not be generalized to patterns of news consumption that might be found in a broader society.

The estimation of hours per day for using traditional media and the Web was resulting from the answers to open-ended questions, concerning days per week or hours per day respectively. Only 20 percent of respondents were more likely to use the Web for news than to use newspapers or television. On the other hand, 60 percent of respondents were keener to rely on television news than the Web.

The average time spent watching news programs on television was equal to approximately 35 minutes per day, and the mean time dedicated to entertainment programs reached almost 100 minutes per day. Two approximations of the time spent on the Web were used: estimation of days per week on which the Web and number of days per week on which the Web was used for surveillance.

Outcomes of the conducted analysis indicate that the use of the Web for watching news supplements rather than substitutes for the use of traditional news media in the analysed community. It also appeared that traditional news media do not seem to compete directly with each other or with the Web for news - the more time the respondents spent with one medium, the more time they spent with the other. The obtained results also show that likelihood of using the Web for watching news has no statistically significant relationship with watching television news programs.

Another approach to study usage of online news compared with news use via traditional media was based on niche theory and the theory of uses and gratifications. Survey presented in [5] was based on data which were collected in a telephone survey with 211 respondents in the Ohio (USA) metropolitan area. The results indicate that the Internet has a competitive displacement effect on traditional media in the daily news domain with the largest displacements happening for television and newspapers.

The findings also show that there is a relatively high degree of overlap between the niches of the Internet and the traditional media. Moreover, the results suggest that the Internet satisfies more needs than any of the traditional media.

Surveys were also done on the relationship between the Internet use and the individual-level social activity. In work presented in [3] the authors apply a motivational perspective to differentiate among types of Internet use when examining the factors predicting such aspects as civic engagement. The predictive capabilities of new media use are then analyzed relatively to key demographic, contextual and traditional media use variables

Studies were done to verify if “digital divide” in Internet connectivity is reflected in Internet usage [1]. For this purpose, a survey of over 18 thousand American citizens was performed. The main finding was that educated people with relatively high income were more likely to have adopted the Internet earlier in time. However, conditional on adoption, low-income, less-educated people spend more time online. The obtained verification implies that this pattern is best explained by differences in the preferences concerning of leisure time.

In methodology adopted in a survey made by Reuters Institute [7] the collected data were referred to user specific user characteristics, such as age, gender and region. The sample collected is reflective of the population that has access to the Internet. The survey was focused on news utilization; consequently the subjects who declared they had not consumed any news in the past month were filtered out.

An online questionnaire was elaborated in order to cover significant aspects of news consumption – therefore it may under-represent the consumption habits of people who are not online, usually older and with limited education. Additionally, a number of face-to-face focus groups were held in a few European Countries and in the USA for investigating issues related to news consumption.

The main outcomes show that approximately half of respondents use social media as a source of news. Women and young people are predominant in this group.

In addition to online access, most users also continue to access news via TV and radio but the extend of this is significantly affected by age. For every group under 45, online news is now more important than television news. Television news still remains most important for older people however, overall usage of this medium continues to decline. For young people (under 24) social media (28 percent) become more popular than TV (24 percent). There are quite significant differences between some countries regarding news sources: for instance in Ireland, almost 40 percent of citizens start their day with radio news and countries like Finland and Japan significant minority still read printed newspapers in the morning. TV usage ranges from 16 percent in Ireland, up to 51 percent in Japan.

As concerns devices, usage of smartphones for news watching is firmly increasing, especially among young people, reaching over half of respondents. Computer use is declining, while tablet usage seems to be saturating, falling back in some countries like the UK and Japan.

Findings resulting from data across all considered countries show that news articles are still the most consumed type of news online content (59 percent) while online news videos were preferred by 20 percent of users on average (22 percent in France, 27 percent in Spain and 28 percent in Poland).

The above presented studies concerned mainly the patterns of the Internet and traditional news media usage in general. These studies do not refer directly to multilingual help system but give additional information, which can be helpful for

adjusting functionalities and features of multilingual help system to the news usage patterns.

There are not many sources regarding user requirements with respect to video summarisation and machine translation. The analysis presented in this paper focuses specifically on such requirements regarding functionalities combining both: automatic video summarisation and translation of news related content.

### 3 User and Services Requirements

In this section, the authors describe the questionnaire designed to extract information for the users and the subsequent statistical analysis.

#### 3.1 Questionnaire Design

To define user and AMIS service requirements, the authors designed a questionnaire to extract information from users related to their habits to consume news, either through the TV or the Internet. In tables 1, 2, 3 and 4, the parts of designed questionnaire can be seen. As it can be seen the items proposed in the questionnaire are oriented to characterize the user profile taking into account socio-demographic data (such as: educational level and languages that one speaks), the use of sources of information, the preferences of offered functionalities, and human interfaces.

Section A and B (tables 1 and 2) of the questionnaire ask the user about descriptive information, however, section 3 (table 3) allows the user to rank and order the proposed system functionalities. Section D questions are oriented to extract GUI user preferences to guarantee the system usability and accessibility in the future design.

**Table 1** Questionnaire Section A

<b>A. Respondent Information</b>	
A.1 Age	
	Below 18 / 18 – 25 / 26 – 35 / 36 – 45 / 46 – 55 / 56 – 65 / above 65
A.2 Gender	
	Male / Female
A.3 Please indicate your mother tongue language	
A.4 Language – comprehension skills sufficient for understanding news	
	Arabic / English / French / German / Italian / Polish / Spanish / other
A.5 How many times a year do you travel abroad?	
A.6 State your highest completed education level	
	Elementary/High School/Bachelor/Master/Doctoral
A.7 Disability	

**Table 2** Questionnaire Section B

<b>B.Usage of Sources of information</b>
B.1 Frequency of TV news usage
B.2 How long, on average, is your TV news watching session
B.3 Frequency of Internet news usage
B.4 How long, on average, is your internet news watching session
B.5 Would you prefer to Choose TV stations yourself / Rely on suggestions
B.6 Do you ever watch video whose language you do not understand?
B.7 Would you watch such video more frequently if it was translated?
B.8 Which subjects are you interested in: Local news / National news / International news / Business / Sports / Life Style and Culture / Technologies and Science
B.9 Which form of programs you prefer Discussion of experts / News supported with videos / Films presenting a subject

**Table 3** Questionnaire section C

<b>C. Functionalities</b>
Please rank the functionalities
- Watching summaries of a longer content
- Searching for subjects/keywords within a given content
- Automatic translation of foreign language
- Comparison of two different audio-video materials on the same subject
- Interaction with social media
- Further exploring of a given subject

**Table 4** Questionnaire Section D

<b>D. Human interface</b>
D.1 How would you like to have the translations Voice translation over original audio/Text displayed on the original content
D.2 Which equipment would you prefer to use with the proposed functionalities Computer/Tablet/Smartphone/Equipment attached to the TV/Other
D.3 Would you prefer Dedicated application/web browser/other
D.4 Would you prefer to Services profiled/dedicated to you/Service as anonymous user

In table 4, the items D.3 and D.4 are specially oriented to decide the hardware and software requirements for AMIS system.

### 3.2 Statistical Analysis

In the statistical analysis performed, the normality of the data by Levene test [4] for homogeneity of variances, and the Kolmogorov-Smirnov test [6] has been previously checked. The previous tests have proved that the data obtained through

the questionnaires don't belong to a known distribution. Thus, non-parametric methods have been chosen, in particular, Kruskal Wallis test [2].

In the following subsections, the results obtained from the users are presented, with respect to gender and age, mother tongue as well as education level.

### 3.2.1 Socio-demographic Analysis

In this study 174 people have answered to the previously described questionnaire, mainly all of them are between 18 and 65 being 60,90% males and 39,1% females (see table 5). Many of them have at least bachelor degree.

#### (A) Gender and Age

As it can be seen in table 6 the participants were divided in seven age ranges being most of them (81,6%) under 45. In fact, the core of the group is between 18 and 45 (63,3%). With Kruskal Wallis test [2] (this non-parametric test check if input data belong to the same population), significant differences are checked between the two groups (male and female) divided by gender independently of the age.

**Table 5** Gender data

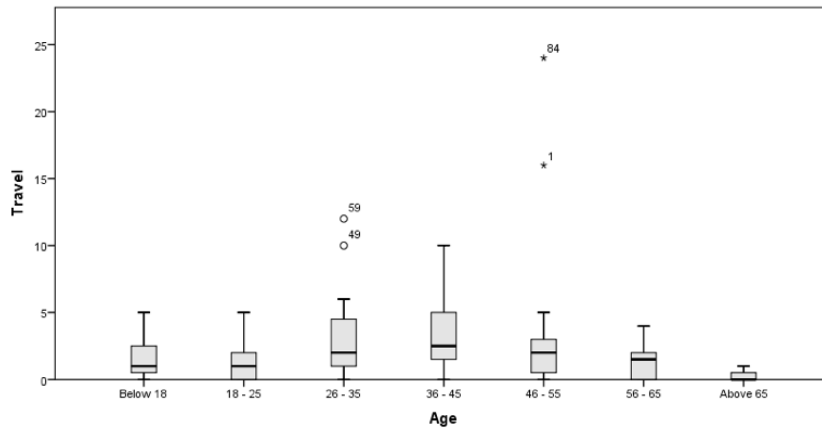
	Frequency	Percent
Male	106	60.9
Female	68	39.1
Total	174	100

**Table 6** Age data

	Frequency	Percent
Below 18	32	18.4
18 - 25	61	35.1
26 - 35	28	16.1
36 - 45	21	12.1
46 - 55	16	9.2
56 - 65	11	6.3
Above 65	5	2.9
Total	174	100

A significant difference has been found between the 7 groups studying the relation between the number of trips abroad that they perform per year  $p = 0.002$ . The p-value is defined as the probability of obtaining a result equal to or "more extreme" than what was actually observed, when the null hypothesis is true. In this case, the null hypothesis is that a relation between age and number of travels abroad does not exist. The following Fig. 1 shows the box plot for this variable:





**Fig. 1** Age box plot. Subjects have been divided into 7 groups according to age. The number of travels to other countries show significance differences ( $p=0.002$ ). Circles in the chart show outliers subjects.

As it can be seen in Fig. 1 which shows a box plot (it is a diagram which shows the median, the first and third quartile, and whiskers. These whiskers indicate the highest and the lowest value respect to the inter quartile range) with the information of the age ranges in which most people travel abroad (the three central ranges: 26-35, 36-45, 46-55).

In Fig. 1, 4 outliers can be observed. They are found in the ranges 26 -35 and 46 - 55. They are participants who travel more than the average.

#### (B) Mother Tongue

In relation to mother tongue, there is a significant difference regarding the number of additional languages that the participant speaks and understands in an acceptable way ( $p = 0.001$ ). Usually the users speak/understand correctly French, Spanish and Polish (due to the origin country of the authors).

The statistics (table 7) show that people whose mother tongue is French or Arabic speak at least one extra language. However, in the case of people from Spain and Poland, this number increases to 2. On average, all the participants speak 2 extra languages.

**Table 7** Mother tongue data

	Frequency	Percent	Average of extra languages
Arabic	4	2.3	1.50
French	39	22.4	1.64
Spanish	31	17.8	2.61
Polish	93	53.4	2.09
German	3	1.7	3
English	1	0.6	1
Italian	1	0.6	4
Others	2	1.2	1.5
Total	174	100	2.09

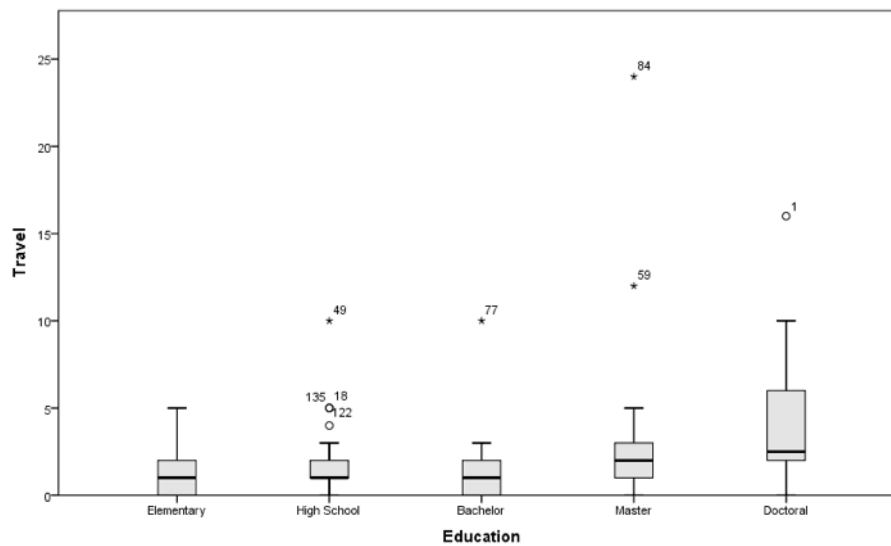
## (C) Education Level

In table 8, the educational level of each participant can be seen. More than 50% of the users have at least the Bachelor Degree.

Regarding the educational level of the participants, using the Kruskal Wallis test, we have found significant differences in the number of trips abroad that they perform a year ( $p = 0.000$ ). There is a significant difference in the number of travels abroad between subjects divided according to education. The following Fig. 2 shows the box plot for this variable.

**Table 8** Education level data

	Frequency	Percent
Elementary	28	16.1
High school	53	30.5
Bachelor	30	17.2
Master	40	23
Doctoral	23	13.2
<b>Total</b>	<b>174</b>	<b>100</b>



**Fig. 2** Education box plot. Education box plot. Subjects have been divided according to education. Circles in the chart show outliers subjects.

In Fig. 2, several outliers can be observed. They are participants who travel more than the average.

### 3.3 Usage of Sources of News Information on TV and Internet

Participants have been divided according to age ranges specified previously. After making the Kruskal Wallis, significant differences in how often the people watch the news on TV ( $p = 0.015$ ) were observed. Dividing participants questionnaires by gender, we also found significant differences in how often they watch the news on TV ( $p = 0.024$ ).

In relation to mother tongue, there is a significant difference in how often the people watch the news on TV per day ( $p = 0.000$ ), the minutes they spend watching news on television ( $p = 0.000$ ), and the number of times they review the news on Internet ( $p = 0.011$ ).

### 3.4 Functionalities of the Proposed System

Preferences of the respondents concerning the functionalities of the proposed system are as follows:

1. Watching summaries of longer content - 32,18 percent
2. Automatic translation of foreign language - 25,29 percent
3. Searching for subject/keywords within a given content - 13,79 percent
4. Interaction with social media (rating and sharing the content) - 11,49 percent
5. Further exploring a given subject - 10,34 percent
6. Comparison of two different audio-video materials on the same subject - 6,90 percent

The table 9 shows the chosen functionalities for each age range.

**Table 9** Functionalities selection for each age range

<b>Funct.</b>	<b>Below 18</b>	<b>18 - 25</b>	<b>26 - 35</b>	<b>36 - 45</b>	<b>46 - 55</b>	<b>56 - 65</b>	<b>Above 65</b>
1	6,32%	10,92%	5,75%	4,60%	2,30%	1,15%	1,15%
2	5,17%	8,05%	4,02%	4,02%	1,72%	1,72%	0,57%
3	2,87%	3,45%	4,02%	1,15%	1,72%	0,57%	0,00%
4	1,15%	4,60%	1,15%	0,57%	1,72%	1,15%	1,15%
5	1,15%	5,17%	0,57%	0,57%	1,15%	1,72%	0,00%
6	1,72%	2,87%	0,57%	1,15%	0,57%	0,00%	0,00%
<b>Total</b>	<b>18,39%</b>	<b>35,06%</b>	<b>16,09%</b>	<b>12,07%</b>	<b>9,20%</b>	<b>6,32%</b>	<b>2,87%</b>

Observing these results, the two first items are the best rated, and it is completely aligned to the objectives of the AMIS project. All the participants, independently of the age range prefer when the system includes functionalities 1: "Watching summaries of longer content", and 2: "Automatic translation of foreign language" (except people above 65).

### 3.5 GUI Preferences

In this subsection, the answers of section D questionnaire are analyzed, mainly, D3 and D4 items.

Taking into account D3 answers, the participants have been divided according to the preferred technology (dedicated application or web browser application) there is a significant difference in number of news that participants review ( $p = 0.008$ ) and number of minutes spend looking news on the Internet ( $p = 0.012$ )

In the case of D4 no significant differences have been found.

The table 10 shows the preferred technology for each age range.

**Table 10** Preferred technology for each age range

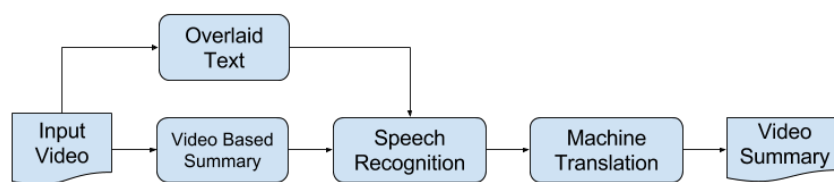
	Below 18	18 - 25	26 - 35	36 - 45	46 - 55	56 - 65	Above 65
<b>Dedicated application</b>	8,05%	13,79%	6,90%	4,60%	3,45%	2,87%	1,72%
<b>Web application</b>	6,32%	19,54%	8,05%	6,90%	5,75%	1,72%	0,00%
<b>Others</b>	4,02%	1,72%	1,15%	0,57%	0,00%	1,72%	1,15%

In the questionnaire, people between 18 and 55 have chosen a web application design of the system.

#### 4 Overall Architectural Recommendations

The analysis of the questionnaires shows, that the functionalities of summarization of the content and automated translation are considered most important for the users. These findings had influence on our approach to the architecture design.

The proposed architecture will be at the same time effective in delivering summarized and translated content and will be fairly easy to implement and integrate. This approach will allow to analyze future research and engineering challenges in order to adapt the second phase architecture. We have also agreed that at first the system will be lacking some of the planned non-critical features such as full social networks integration and sentiment analysis as those are not chosen as most important by our respondents. Proposed phase one architecture is depicted in Fig. 3 with the main pipeline denoted.



**Fig. 3** First phase architecture

We start with a full video file which is summarized to desired length. The audio of the summarized video file is extracted and passed to the speech recognizer. An up-to-date speech recognition toolkit, such as KALDI, will be used. The results of the speech recognition are translated using the machine translation algorithms. For machine translation The system will use the state-of-the art Moses system along with the Giza++ toolkit, as well as the IRSTLM language modeling toolkit, as proposed in the Moses' scripts. The summarized video file with added captions provided by the machine translation is presented to the user. At the same time video summarization, speech recognition and machine translation are supported by information provided by the optical character recognition system that is analyzing the overlaid video text.

In the second phase we plan to implement integration with social media and sentiment analysis to cover other functionalities planned for the system. When moving to the second phase we also plan to enhance functionalities and algorithms used in the first phase in order to deliver higher quality summarizations and translations.

## 5 Summary on Impact of Expected Further Results

The paper presents results of the survey made in order to specify profiles of potential users, their habits concerning usage of information as well as their preferences regarding functionality of the multilingual help system supporting the understanding of broadcasting news in foreign language.

Over 170 persons with different cultural and social background were subjects to the analysis. The results of the survey will allow to fine-tune the features of the system.

As a result AMIS will greatly minimize difficulties in the exchange of information, regardless of language differences, which nowadays is extremely important.

Its use concerns mainly social and cultural domains, as it allows one to explore another side of specific event. This will allow users to understand the comments, as well as the culture of other nations.

At the scientific level the project concerns synergies and new research linking summarizing video, audio and text materials, ASR, machine translation and opinion exploration.

Evaluation of project outcomes, multilingual information and opinions, due to its proximity to human experience, will be primarily done subjectively. Besides, some automatic evaluation methodologies are planned to be applied, at least for some of the elements of the proposed solution.

## References

1. Avi Goldfarb, J.P.: Internet adoption and usage patterns are different: Implications for the digital divide. *Information Economics and Policy* **20**(1), 2–15 (2008)
2. Breslow, N.: A generalized kruskal-wallis test for comparing k samples subject to unequal patterns of censorship. *Biometrika* **57**(3), 579–594 (1970)
3. Dhavan V. Shah Nojin Kwak, R.L.H.: Connecting and disconnecting with civic life: Patterns of internet use and the production of social capital. *Political Communication* **18**(2), 141–162 (2001)
4. Glass, G.V.: Testing homogeneity of variances. *American Educational Research Journal* **3**(3), 187–190 (1966)

- 
5. John Dimmick, Y.C.Z.L.: Competition between the internet and traditional news media: The gratification-opportunities niche dimension. *Journal of Media Economics* pp. 19–33 (2009)
  6. Lilliefors, H.W.: On the kolmogorov-smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association* **62**(318), 399–402 (1967)
  7. Nic Newman with Richard Fletcher, D.A.L.L., Nielsen, R.K.: Reuters institute digital news report 2016 (2016)
  8. Scott L. Althaus, D.T.: Patterns of internet and traditional news media use in a networked community. *Political Communication* **17** (2000)