FUTUREWEB
deep learning on big data

PN-III-P1-1.2-PCCDI-2017-0800
Modelarea empirică și dezvoltarea experimentală a instrumentelor asociate
tehnologiilor emergente din domeniul rețelelor sociale online - **86PCCDI/2018**

Summary of sub-project 5 activities: Experimental development of a Romanian-specific ontology and testing the effects of semantic web-based searches at the level of online social networks" (Semantic Media)

-Summary 2018-

Activity 5-1: Identify the role of using semantic technologies at the level of online social networks

The virtual environment consists of a huge "network" of information resources, created for the purpose of dissemination of its products by users, most of the information, designed to respond to user demand, information, which most often is not hierarchical or structured. A Google search returns about 25% of all user-relevant results, and very often some searches may not return any results, even if there are sites relevant to that search.

Thus, there was an acute need for accessible identification of data, as well as a need to filter, interpret and synthesize information, for the benefit of users, with the help of programs and new technologies developed. The solution is to provide users with a way, based on current XML technologies, to process information by the computer, through a scenario called the Semantic Web.

The Semantic Web can be perceived as a system that allows computers to "understand" and respond to complex user requests, based on meaning and significance. At the same time, it is considered an integrator of content from contexts, systems, and information applications.

The Semantic Web allows data, located anywhere on the web, to be accessible and understood by both humans and computers. This can be done by adding extensions, called meta data, to existing documents. This allows data to be processed automatically by computers, according to the same principles that are manually processed by humans. The Semantic Web is dependent on the existence of specific languages and the multiplication of web pages annotated with meta data. Basically, when we upload a certain file on the internet, it will be attached a meta data that will "know" to transmit to

PN-III-P1-1.2-PCCDI-2017-0800
Modelarea empirică și dezvoltarea experimentală a instrumentelor asociate
tehnologiilor emergente din domeniul rețelelor sociale online - **86PCCDI/2018**

a possible existing search engine, details about the content and meaning of the entire article.

Due to the large number of potentially useful data created on a variety of social networks, the direct use of social networks is impractical to extract useful information. Therefore, advanced filtration techniques are required. Searching for information on social media is based on identifying words and matching them.

The Semantic Web promises a cure. The idea of a Semantic Web involves a move from unstructured sites (for example, without or only with rare meta data, understandable by the computer) to structured ones, which cannot be understood only by people. Semantic vocabulary is based on concepts that are defined in ontologies.

Ontologies provide a vocabulary of terms in a particular field, which is needed to detail the meaning of annotations added to websites. Ontologies can unify and normalize data from different resources, e.g. social media, syntactic and semantic and associate them with knowledge in the field of interest. Moreover, the ongoing research has shown that it is important to have a common understanding of concepts within and between areas to avoid misunderstandings.

To be used by the semantic web, ontologies must be used by computers and thus must be expressed in languages that computers and people can easily understand. Thus, the main languages for defining ontologies are based on XML - a language that is very easily interpreted by computers.

A common feature of ontologies is that they are formal. Due to this feature, the construction of an ontology is particularly complicated, because it must contain many correct and especially non-contradictory logical relationships.

In order to be able to develop an ontology specific to the Romanian language for social networks, two qualitative researches were used among the stakeholders involved (organizations and individual users) and a quantitative research among the users of the social networks individuals.

Thus, following 30 in-depth interviews and based on the results of qualitative research among the representatives of the organizations studied, the following conclusions were drawn:

- the study of the perception of organizations towards online social networks highlighted the fact that the representatives of these organizations associate

PN-III-P1-1.2-PCCDI-2017-0800
Modelarea empirică și dezvoltarea experimentală a instrumentelor asociate
tehnologiilor emergente din domeniul rețelelor sociale online - **86PCCDI/2018**

the concept of social network with Facebook, but also with words such as communication, connection, and information.

- the analysed organizations stated that they have accounts in online social networks, mainly on Facebook, Twitter, Instagram, YouTube and LinkedIn, having at least two of these two active accounts, the frequency of use being daily or weekly

- the most used accounts for professional purposes are Facebook, LinkedIn and YouTube, the searches being performed by keywords / phrases that are related to the field of activity of the studied organizations

- organizations use, for a fee, services offered directly by online social networks, such as Facebook Insights and less Twitter Analytics, Pinterest Analytics, or others

- no organization has indicated the use of other analysis and monitoring services of online social networks, other than those offered directly by social networks, such as Hootsuite, SAP Social Engagement Cloud, etc.

- half of the organizations analyse and monitor social networks and for products and services offered by other organizations with which they are in competition

- the concept of Semantic Web is not known among the studied organizations, being considered a term with a remarkably high degree of novelty

- at the personal or organizational level, Semantic Web technologies are not used at all, the reason being the lack of information on the practical usefulness of semantic technologies and, possibly, the lack of an identified need for their use for organizational purposes

- the most common types of functionalities considered to be implemented with the help of Semantic Web-specific technologies within the organization are those that allow the analysis of users' opinions on the characteristics of a product / service, semantic search / filtering and advanced understanding of message content. these being considered the most useful functionalities

- analysis of user relationships, analysis that can identify influencing users who promote a good image of the organization or those who promote a negative image of the organization, did not record responses from organizations

PN-III-P1-1.2-PCCDI-2017-0800
Modelarea empirică și dezvoltarea experimentală a instrumentelor asociate
tehnologiilor emergente din domeniul rețelelor sociale online - **86PCCDI/2018**

studied, which may suggest a misunderstanding of the usefulness of that functionality

- organizations suggest a number of utilities of semantic technologies, such as the rapid transmission of data on the organization's products / services, the improvement of products / services offered to customers based on advanced searches, the rapid search for information relevant to the organization's activity to take decisions, obtaining a clear picture of the opinions of online social network users, developing new tools or associated services based on the needs expressed in the online environment and others.

Following a qualitative research among individual users of social networks, through the 30 in-depth interviews the following conclusions can be drawn:

- all respondents have a social media account, mainly Facebook and Instagram, accounts that are accessed daily for clearly defined purposes: communication, socialization, and information

- most respondents associate the concept of social network with Facebook or keywords such as: interaction, communication, and socialization

- the most accessed services for analysing and monitoring online social networks consider the use of services offered directly by online social networks, more than half of the respondents indicating that they use Facebook Insights and very few Pinterest Analytics or Twitter Analytics

- no respondent uses online social network analysis and monitoring services, other than those offered directly by social networks, such as Hootsuite, SAP Social Engagement Cloud

- analysis and monitoring services of social networks are performed for personal purposes, free of charge by all respondents

- only 3 respondents heard about the semantic Web concept, of which only one is familiar with what semantic ontology means and types of functionalities specific to the semantic web

- respondents faced several problems related to the difficulty of finding answers to more complex questions and the existence of lexical differences, in the sense that the same meaning can be expressed in different ways.

PN-III-P1-1.2-PCCDI-2017-0800
Modelarea empirică și dezvoltarea experimentală a instrumentelor asociate
tehnologiilor emergente din domeniul rețelelor sociale online - **86PCCDI/2018**

- problems have been reported regarding the ambiguity of the result, in the sense that the same word or sentence may have different meanings and the multilingualism generated by the expression in several languages

- as types of functionalities of the semantic web considered useful were indicated the semantic search that takes into account the different meanings of a word and the semantic search that takes into account the synonyms of the words used, respectively of words with more specific meaning

- analysis of users' opinions on the characteristics of a product or service in online social networks were of interest to respondents to be implemented using semantic ontologies

- the opinions of other users of online networks are important insofar as they concern issues related to socio-political, domestic, or international events and those related to the analysis of other users' reactions to catastrophic events.

- the existence of a rather high degree of ignorance of the usefulness of the different types of functionalities specific to the semantic web was noticed, but no cognitive barrier was registered regarding the subsequent promotion of this concept.

Based on the qualitative research, a quantitative survey research was designed, by distributing an online questionnaire on the platform www.esurvayspro.com.

Following the collection of data and their analysis, the following conclusions can be drawn:

- only 30.43% have heard of this concept, and of those who have heard, 73.91% are familiar with Semantic Web technologies to a small extent, 26.09% on average

- although a high proportion of respondents have heard of these technologies, they do not know the advantages offered by semantic technologies in social networks

- most of the respondents 63.47% mention as the first social network that comes to mind Facebook, followed by Instagram with 24.35%, and quite long-distance LinkedIn with 5.22%, YouTube with 4.35% and Twitter with 2.61%

- areas of interest of respondents in social networks, they can be grouped into the following areas: fashion and beauty services, music and art, fashion, business (entrepreneurship, marketing, financial education), personal development, leisure and relaxation, current affairs political and social, career opportunities

5

PN-III-P1-1.2-PCCDI-2017-0800
Modelarea empirică și dezvoltarea experimentală a instrumentelor asociate
tehnologiilor emergente din domeniul rețelelor sociale online - **86PCCDI/2018**

- Regarding the most common words used in social media searches, respondents' answers show their interest in friends, restaurants, weather, quotes, fun, news, events, brand names, marketing, fashion, photos, fashion, etc.

- Regarding the problems faced by users in social media searches, the first element mentioned by 59.8% of respondents was "semantic difficulties" (same meaning expressed in different ways), followed by ambiguity (same word/ sentence has been differently understood), mentioned by 55.3% of the respondents, multilingualism (expression in several languages of the searched words), mentioned by 50% and difficulty in finding answers to more complex questions, mentioned by 10.71% of those interviewed

- the advantages (functionalities) offered by semantic technologies (semantic ontologies) are highly appreciated by users, in the first place being the analysis of other users 'opinions on the characteristics of a product, followed by analysis of other users' opinions on companies, organizations and emotions of users regarding the elements of the presentation

The creation of ontologies in Romanian for social networks becomes imperative, because it leads to solving major problems of the semantic internet: classification, indexing and more accurate retrieval of knowledge contained in social networks, development of multilingual interfaces to specific web applications semantically, the creation of virtual communities free of language barriers, etc.