

Chapter 1

Computer Analysis of Texts in Social Networks, Its Method and Tools: State-of-the-Art Review

Pavel P. Makagonov

*The Russian Presidential Academy of National Economy and Public
Administration, Russia*

Alexander Troussov

*The Russian Presidential Academy of National Economy and Public
Administration, Russia*

ABSTRACT

Texts in virtual social networks differ cardinally from those of reviewed and edited publications, as being in fact the materials of non-moderated chat dialog having all the syntactic features, and in many cases, they are hypertexts with, as analysis subjects, quite relative boundaries. Instead of texts the virtual discourse, an object of new type is analyzed. In these conditions, the linguistic analysis is transformed into preliminary linguistic processing the texts and analysis of texts into the analysis of networks participants' interactions. Weak sides of linguistic processing both primary and pervasive are shown concerning the texts in social networks. Many problems appear when analyzing the texts as non-structured big data. A changing of used linguistic tools is linked with the reorientation onto new research customers: companies seeing the business interest in that, state structures, and even onto criminal structures using result of analysis for composing the texts aimed for manipulation of the consciousness.

DOI: 10.4018/978-1-5225-5586-5.ch001

INTRODUCTION

Texts found in virtual social networks fundamentally differ from those found in publications that have been subjected to editing and reviewing processes. Social network texts are texts that are retrieved from non-moderated chat-dialogue and that have specific syntactic structure. In many cases, they are hypertexts that have notional boundaries when taken as objects of research analysis. When researchers shift from the analysis of professionally created texts to the analysis of texts that are created by Internet users in a variety of virtual social networks, they face challenges associated with both research object and the shift in the pragmatic character of (research customer) expectations related to the research results. Texts created in virtual social networks have their own specific characteristics in terms of authorship, stimuli and agendas for altering vocabulary as well as of the ways they convey meaning, employ specific grammar and select topics for discussion.

The tools of linguistic analysis were developed a long time ago with the primary focus on the lexical analysis of literary fiction, authentication of the texts authorship and, later, creation of computer-assisted translation and automated text referencing and indexing. With the rapid development of mass media, new goals for linguistic analysis emerged: processing large volumes of text, creating and indexing large e-libraries, speech recognition, information searching, computer-assisted document classification and referencing, computer managed linguistic processing, creating online dictionaries, constructing linguistic processors to ensure users' communicating with automated intellectual systems, and extracting fractographic information from informal texts, to name but a few. A significant change in linguistic analysis, its tools and objectives was stimulated by the sharp increase in the volume of information found on the Internet (proliferation of news channels, e-libraries, non-structured texts stores, anti-plagiarism software development, etc.) All this, coupled with the use of computer technologies and networks, led to the change in the linguistic analysis toolkit. Pre-processing services became available and the field of corpus linguistics flourished. When applied to the above-mentioned tasks, the linguistic analysis toolkit relies on a systematically developed theory and is complemented by specific software.

The development of social networks prompted the need to revise the methods of computer linguistic analysis given the new object of research and its new tasks. Before the appearance of the Web, computer linguistic research related to the area of corpus studies could have focused, for example, on the work of a specific writer. A social network on the Internet (or a virtual context) comprises many users, who, united by a common goal or interest, create their content. Specific "thematic" online forums, that have recently experienced rapid development, likewise belong to this category. In his "Socio-Dynamics of Culture," Moles (2008) divided society into

26 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/computer-analysis-of-texts-in-social-networks-its-method-and-tools/208378

Related Content

Relational Dynamics: The Flood Solution – Sustainable Energy and Rural Tourism

David Stefan Bathory (2017). *Applied Behavioral Economics Research and Trends* (pp. 128-151).

www.irma-international.org/chapter/relational-dynamics/173649

Channel Conflict and Management of O2O Network Marketing Model Under E-Commerce Exploration of Ideas

Rafia Sber (2022). *International Journal of Circular Economy and Waste Management* (pp. 1-4).

www.irma-international.org/article/channel-conflict-and-management-of-o2o-network-marketing-model-under-e-commerce-exploration-of-ideas/312227

Clusters for National Growth: Belt of Prosperous Nepal, Projects Along the North-South Economic Corridor

Madhukar Thapa (2022). *Handbook of Research on Special Economic Zones as Regional Development Enablers* (pp. 135-169).

www.irma-international.org/chapter/clusters-for-national-growth/291135

Online to Offline-based e-waste "Internet + Recycling" pattern building:

Online to Offline-based e-waste

(2022). *International Journal of Circular Economy and Waste Management* (pp. 0-0).

www.irma-international.org/article//311052

New Method New Estimates

(2014). *Bridging the Gap Between Growth Theory and Policy in Asia: An Extension of the Solow Growth Model* (pp. 185-216).

www.irma-international.org/chapter/new-method-new-estimates/104604