

GREY LITERATURE BETWEEN TRADITION AND INNOVATION: IS THERE A CONTINUUM?

Gabriella Pardelli, Sara Goggi, Manuela Sassi
Istituto di Linguistica Computazionale, "Antonio Zampolli"
Consiglio Nazionale delle Ricerche (CNR)
Via G. Moruzzi 1, 56124 Pisa, Italia
[gabriella.pardelli, sara.goggi, manuela.sassi]@ilc.cnr.it

Abstract

This study wants to explore new ways of social media communication for Grey Literature. In particular it describes the role of social media in relation with traditional channels and how social media applications can be used for Grey.

Keywords: Grey Literature, Communication networks, Knowledge networking, Knowledge exchange.

1. Introduction

From the Vth century, when monasteries started to be built in Western Europe, the transcription of ancient literary works was made in one of the abbey rooms called 'scriptorium'. This method for the replication of knowledge aimed at retrieval and conservation rather than at divulgation: there was a selection of the works to bequeath and preference was given to Latin Classics, religious works and some heathen works as well.

This Medieval way of knowledge was preservative rather than popular and was addressed to a niche of a few users: its main merit was to pass on knowledge from age to age and transmit it to our time Information Society (IS). Information Society is a computer-based society which spreads knowledge, meets on social networks, establishes realities such as Wikipedia and Facebook thus reducing the world into a global village.

Information currently means knowledge and becomes a new form of wealth, since productivity and welfare are influenced by the management and the transmission of information. The Medieval "scriptorium" is somehow rebuilt on the web with a new transcription system called "digitalization":

knowledge is not only created but also disseminated for then being shared. The computer-based culture offers people the opportunity for a cultural growth while technology creates new professions daily and enriches the traditional crafts with new contents.

Nowadays on the web there is a rapid growth of new ways of learning dictated by the massive information exchange and the shared knowledge: “The web is immense, free and available by mouse-click. It contains hundreds of billions of words of text and can be used for all manner of language research” (Kilgarriff and Grefenstette, 2003).

Internet has become a universal repository where users, by means of web-based interfaces, can extract information. But not only: the web makes available to users what is called “social networking”, that is a tool of interactive approach, a collective voice and a new way of communicating. The interactive nature of this tool makes information expand and creates an added value which enlarges knowledge.

Discussion and shared interest for knowledge is the adhesive of social network members; and social networks are a sort of modern literary saloons where the meetings have a “free” and spontaneous nature, participants have a common socio-cultural background and the participation itself is the intrinsic purpose. This said, the importance of social networks lies just in the richness of relations – human, cultural, scientific, economic - which can be developed.

Is the web – alongside with its new ways of social media communication - the modern form of medieval “scriptorium”? How do the social media applications relate to Grey Literature and how can they be employed for research on Grey?

Is there a continuum? Answering to these questions will be the scope of this article.

2. Why Social Media, now?

In January 2010, *The Economist* published a special report on social networking titled “A world of connections”: on page 14 we read: “This democratization of technology is driving the socialisation

of the web and fundamentally changing the way that people interact with one another, as well as with businesses and governments”.

The use of technology becomes democratic, that is available and easily accessible to most people regardless of age, culture, country and economic condition. This is the first time in history that a free communication system increasingly becomes a flow of information through space and time. This process does not duplicate knowledge as it happened in the Medieval ‘scriptorium’ nor store it but is a simple transmission of news and events happened all over the world, a stream of thought which relates all those who are connected.

Nowadays the concept of *Social Networking Sites (SNS)* does not imply the presence of a final user considered as a target because everyone has the same needs and the social network becomes a virtual society, a tribe where the community can find reasons to join. Social networking therefore stands for a tool of interactive approach and means a collective place where people aggregate and communicate: information runs on the web thus enlarging and making knowledge shareable to most citizenship. This is the richness of today’s society, where the rapid increase and the ever stronger need of human relationships run alongside. Google – first brand in this society's world - is its new paradigm: the immaterial.

2.1 Why Social Media for Grey , now?

Over the last decade both Grey Literature and traditional literature underwent the transformation of the communication channels of scientific information: the consequence has been a drastic change of the editorial look which became increasingly headed towards electronic publication on the web. For a long time books, journals and grey material have been shipped to users by ordinary transport links while nowadays information runs on the web and delivery means open access for various types of documents such as publications of digital nature from the origins and publications which have been digitally transformed subsequently.

This material has been stored little by little on ad-hoc academic, industrial and governmental web sites and afterwards, at a international level, common topics have been further assembled on powerful infrastructural web sites (in our field good examples are repositories such as OpenSIGLE and the ACL Anthology). The aim is to concentrate knowledge for then disclose, disseminate and share it with citizens: obviously the purpose is no longer to conserve and preserve knowledge for transmission to a niche of few users as it used to be in the Middle Ages.

These days the interactive nature of social networking allows a very close – and inconceivable till a few years ago - relationship between producers and users of knowledge: the rapid quickening of information delivery causes the gathering of politicians, academics, investors for defining politics, generating knowledge and producing wellness. By means of the social networking sites system (SNSs) – the new added value to production and expansion of knowledge – Grey Literature is involved as well, but how? Groups of Grey Literature scholars with common interests and goals discuss, ponder and compare with new methodologies and new type of information for trying to retrieve/define new paradigms. In this modern XVIIIth century-like literary saloon, culture is created and spread around with the purpose of establishing clusterings of contents and objectives. For sure a new Grey generation is born and the answer to the question “ why social media for Grey, now” is very simple: “It is just moving with the times”.

3. Grey Literature between tradition and innovation:is there a continuum?

“Les concepts font l’objet d’une description systématique qui recens different types d’informations don’t les plus fréquentes sont: le terme utilize pour designer le concept, la definition de concept via celle du terme utilisé, le domaine et le sous- domaines d’employ du terme, les éventuelles variantes (orthographiques ou synonymes de type géographique, par exemple) et les equivalents dans d’autres langue.”(Kister, Jacquy, Gaiffe,2000:27)

The importance of terms – seen as anchors for information retrieval and pointers of time, science and technology - will be discussed in the next paragraph.

3. 1 Social Networking and words for Grey

The linguistic tools developed within the “DylanLab Lab for Computational Models of the Dynamics of Language and Cognition” of the Institute of Computational Linguistics allowed a textual analysis of the abstracts belonging to the 'Social Networking' session of this *Thirteenth International Conference on Grey Literature*.

Table 1. Single Terms

ID	single terms	relevance
1	knowledge	100.0% 
2	networking	95.1% 
3	content	76.0% 
4	library	75.5% 
5	web	71.3% 
6	database	71.3% 
7	literature	71.2% 
8	user	70.8% 
9	information	68.8% 
10	tool	68.6% 
11	communication	67.0% 
12	access	60.4% 
13	community	58.5% 
14	environment	51.9% 
15	research	51.4% 
16	sharing	50.7% 
17	site	50.6% 
18	health	49.1% 
19	representation	45.4% 
20	technology	42.9% 

Table 2. Relevant Multiple Terms

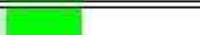







ID	relevant multiple terms (general purpose relevance)	relevance
1	grey literature	100.0% 
2	ability of user	60.4% 
3	new searchable database	60.4% 
4	interest group	52.4% 
5	presence blending information product	42.4% 
6	community of grey literature	40.2% 
7	discoverability of grey literature	39.2% 
8	kind of grey literature	38.6% 
9	interactive communication	38.1% 
10	social networking	38.1% 
11	incremental approach	38.1% 
12	mode of knowledge sharing\creation	38.1% 
13	range of social networking	38.1% 
14	concept from large document	38.1% 
15	area of general medicine	38.1% 
16	many more specialized example	38.1% 
17	communication with much variation	38.1% 
18	tool for volatile knowledge	38.1% 
19	online research interest group	38.1% 
20	realm of human interactivity	38.1% 

Table 3. Domain-Specific Multiple Terms

ID	domain-specific multiple terms (text-dependent relevance)	relevance
1	grey literature	100.0% 
2	social networking	100.0% 
3	ability of user	100.0% 
4	new searchable database	100.0% 
5	interactive communication	100.0% 
6	incremental approach	100.0% 
7	interest group	100.0% 

Although informative contents are usually independent from the technological wrapping, the analysis of the words of the three tables suggests that there should be a connection between inside contents and outside technology: as a matter of fact 12 out of 20 terms of Table 1 refer to the Information and Communication Technology (ICT) domain (*knowledge, networking, web,*

database, user, information, tool, access, communication, sharing, site, technology). The first seven terms have a relevance between 100% and 68,8%.

As far as the relevant multiple terms of Table 2 are concerned, 10 of them are referring to topics proper to Computer Science domain: *ability of user, new searchable database, interest group, presence blending information product, interactive communication, social networking, mode of knowledge sharing/creation, range of social networking, tool for volatile knowledge, online research interest group* and, at last, *realm of human interactivity*, just in reference to social networking. The relevance of the first six terms is between 60,4% and 38,1%.

Finally in Table 3, the domain-specific multiple terms become seven and only the first one, *Grey Literature*, does not refer straight to Computer Science while the other six words can be divided in two subsets: 1) *social networking, interactive communication, incremental approach, interest group*; 2) *ability of user* and *new searchable database*. The four terms of the first set are the perfect summary of the second paragraph of this article while the two terms of the second set are directly connected with the concept of repository and the important ability to consult the stored information for extracting knowledge.

In this scenario the role of information extraction technology becomes increasingly relevant: the priority is to clean up the scene, to sort out linguistic phenomena such as synonymity and homonymy by using domain ontologies. Simply retrieving digital information from academic, industrial and governmental documentation is not enough: nowadays intelligent information retrieval is needed and required.

“Nowadays the need of retrieving the great amount of digital knowledge available on the web is ever more important: the vast majority of this knowledge is conveyed by means of textual material stored in scientific documentary repositories and digital archives. This documentary word preserve inside the wealth of far-off terms belonging to the past which have often fallen out of use as well as

a more recent terminological production derived from the feverish need of coining new terms, a very common trend in certain branches” (Sassi et alii , 2009).

Scientific dissemination in the field of Grey has therefore to be closely connected to the possibility of being easily and quickly retrieved from the web infrastructures.

To conclude, it might be asserted that finding continuity between the old traditional way of producing and sharing Grey Literature and the current technological methods is a difficult task; continuity is rather to be found at the level of contents because Grey Literature will always be Grey Literature, over time and technology.

References

- BANKS M. (2010). Blog Posts and Tweets: The Next Frontier for Grey Literature. In Dominic J. Farace, and Joachim Schöpfel (Eds.), *Grey Literature in Library and Information Studies*. De Gruyter. 217-226.
- BUHL HANS U. (2011). From Revolution to Participation: Social Media and the Democratic Decision-Making Process. *Business & Information Systems Engineering*, Volume 2, Number 4. Online First™, 16 June 2011. DOI 10.1007/s12599-011-0166-4. Gabler Verlag 2011.
- This article is also available in German in print and via <http://www.wirtschaftsinformatik.de>: Buhl HU (2011) Von der Revolution zur Partizipation: Social Media und der demokratische Willensbildungsprozess. *WIRTSCHAFTSINFORMATIK*. doi: 10.1007/s11576-011-0279-x.
- CHIH-LU LIN, HUNG-YU KAO (2010). Blog popularity Mining. Using Social Interconnection, Analysis. *IEEE Internet Computing*, July/August. 41-49
- GUPTA S. THAKUR N. (2010). Semantic Query Optimisation with Ontology Simulation. *Semantic Query Optimisation with Ontology Simulation*, 1(4), 1 - 10.
- KILGARRIFF A., GREFFENSTETTE G. (2003). Introduction to the Special Issue on the Web as Corpus. *Computational Linguistics*, Volume 29, 3, Association for Computational Linguistics, 333-347.
- Kister L., Jacquy E., Gaiffe B. (2008). Réperage de la référence à partir du thesaurus, de la terminologie et de la sémantique lexical. In *Terminologia analisi testuale e documentazione nella città digitale*. AIDA informazioni 1-2/2008. 25-36.
- MAXWELL HARPER F., XIN LI S., CHEN Y., KONSTAN JOSEPH A. (2007). Social Comparisons to Motivate Contributions to an Online Community. *Lecture Notes in Computer Science*, 2007, Volume 4744, Persuasive Technology, Pages 148-159.
- PIEVATOLO MARIA C. (2010). Back to the future: authors, publishers and ideas in a copy-friendly environment. In *Open Scholarly Communities on the Web: Legal, Economic and Social Framework*, Oxford (United Kingdom), 21-22 April 2008. (Unpublished) [Conference Paper].
- ROSEN D., BARNETT GEORGE A., KIM JANG H. (2011). Social networks and online environments: when science and practice co-evolve. *Social Network Analysis and Mining*, 2011, Volume 1, Number 1, Pages 27-42.
- SASSI M., PARDELLI G., BIAGIONI S., CARLESI C., GOGGI S. A Digital Archive of Research Papers in Computer Science. In: *LREC'10 - Seventh International Conference on Language Resources and Evaluation* (Valletta, Malta, 17-23 May 2010). In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odjik, Stelios Piperidis, Mike Rosner, Daniel Tapias (eds.), Proceedings. Paris, European Language Resources Association ELRA, 2010. Pages 1245 - 1248.
- SILLANPÄÄ N., ÄLLI S., ÖVERMARK T. (2010). Easy-to-Use Social Network Service, *Lecture Notes in Computer Science*, 2010, Volume 6179, Computers Helping People with Special Needs, Pages 544-549.
- SCOTT J. (2011). Social network analysis: developments, advances, and prospects. *Social Network Analysis and Mining*, 2011, Volume 1, Number 1, Pages 21-26.
- SØRENSEN ANNE T., SHKLOVSKI THE I., (2011) Hugging Team: The Role of Technology in Business Networking Practices. In Susanne Bødker, Niels Olof Bouvin, Volker Wulf, Luigina Ciolfi and Wayne Lutters (eds.), *ECSCW*

- 2011, *Proceedings of the 12th European Conference on Computer Supported Cooperative Work*, 24-28 September 2011, Aarhus Denmark. Springer-Verlag, London . 333-352.
- SØRENSEN LENE T., SKOUBY KNUD E. (2009). Requirements on Next Generation Social Networking—A User’s Perspective. *Wireless Personal Communications* ,2009, Volume 51, Number 4, Pages 811-823.
- THE ECONOMIST, (January 30th 2010) .A World of Connections: A Special Report on Social Networking’.
- WODZICKI K., SCHWÄMMLEIN E., CRESS U. (2009). How to Get Proper Profiles? A Psychological Perspective on Social Networking Sites. *Lecture Notes in Computer Science*, 2009, Volume 5794, Learning in the Synergy of Multiple Disciplines, Pages 338-343.
- OpenSIGLE Repository, System for Information on Grey Literature in Europe:
<http://www.greynet.org/opensiglerepository.html>
- ACL Anthology, A Digital Archive of Research Papers in Computational Linguistics:
<http://aclweb.org/anthology-new/index.html>