Awareness and empowerment in document production and distribution as a "must" for open access: experiences from the "Nancy style" to guarantee quality

Paola De Castro (1), Sandra Salinetti (1), Marcus Banks (2)

(1) Publishing Activities, Istituto Superiore di Sanità, Rome (Italy)
(2) New York University Medical Center, New York (USA)
e-mail: paola.decastro@iss.it

Background

The debate on Grey Literature (GL) has now a very long tradition going from uncertainty and confusion, in the last century, to new certainties and appraisal in the open access era. This implies the general acquisition of the awareness of the importance of GL as a fundamental primary source of information, hence the necessity to empower authors and issuing organization to guarantee quality in both production and diffusion of GL. During the last international conference on GL held in Nancy in 2005, the proposal for the adoption of an *ad hoc* GL production style was discussed and most welcome by the grey community.[1] This led to the creation of the Grey Literature International Steering Committee (GLISC) which issued, in March 2006, the *Guidelines for the production of scientific and technical reports*,[2] informally known as "Nancy style" (freely available from www.glisc.info).
In the creation of the *Guidelines*, suggestions arising from different realities were closely evaluated and it was not always so easy to reach an agreement that would satisfy all requirements. When the first version was released, there was awareness that some missing points should be developed soon, particularly all issues related to digital GL and metadata.

Goal

This paper intends to reflect on the necessity to update the *Guidelines*, one year after the first draft was presented in Nancy, taking into consideration the rapid changes in the information market. We are convinced that the *Guidelines* may represent a useful tool to improve quality of GL, increase the general awareness of its value, and empower authors and issuing organizations to allow a correct autonomous production of documents to be spread and exploited at best through the Internet. For the same reason, the best strategies to promote the adoption of the *Guidelines* will be investigated: involving GL authors and institutions, making translations available in different languages, establishing useful contacts with key persons in the information arena.

Actions developed

With the objective to create major awareness on quality issues in GL creators, users, and policy makers, the following actions were developed:
– analysis of GL in the changing context;
– reflections on the development of the "Nancy style";
– comparison between "Nancy style" and ANSI/NISO Z39.18;
– promotion of the "Nancy style".

---

[1] De Castro P, Salinetti S. "Uniform Requirements" for grey literature: proposal for the adoption of "Nancy style". *Publishing Research Quarterly* 2006;22(1):12-7.
[2] Grey Literature International Steering Committee. *Guidelines for the production of scientific and technical reports. Version 1.0.* GLISC; 2006. Available at: http://www.glisc.info; last visited Nov. 14, 2006.

Analysis of GL in the changing context

Within the hierarchy of information sources, GL has traditionally occupied a less prestigious status than peer-reviewed papers published in scholarly journals. During the print-only era, this was a valid distinction. Even though GL has always provided access to valuable and unique content, the costs in time and money of obtaining it were formidable. In such a climate, the regularity and reliability of peer-reviewed journals were significant advantages, irrespective of the quality of published papers.

The contemporary ease of electronic distribution has erased a key condition for the traditional hierarchy between grey and white literature. Although this point has not been reached yet, theoretically all scholarly content – be it a data set, a working paper, or a formal paper – could instantly be posted online.

The "open access" movement has focussed on increasing online availability of formal papers, but its general principle is applicable to all scholarly content. Even though distribution of all content has eased dramatically, the firm preference among many scholars for formally peer-reviewed content continues. This is true despite evidence that the consideration of GL enriches many scholarly investigations;[3,4] one of its most valuable characteristics is that it is more likely to report negative results, and to discuss studies that concluded prematurely.[5] Compared to relying only on published articles, also accessing GL offers a much more comprehensive picture of the state of knowledge for any topic.

Because GL is so useful, raising its profile is imperative. One way that the authors of reports can achieve this is to adhere to well-developed production standards. This was the motivation for the development in Europe of the "Nancy style" at the GL7 Conference in December 2005, and for refinement of this style in 2006. In the USA, the *ANSI/NISO Standard Z39.18-2005 Scientific and Technical Reports – Preparation, Presentation, and Preservation*,[6] serves a similar purpose.

The encouragement of quality production standards is an "internal" advance among GL producers. Current developments in the external information landscape also point to an enhanced public awareness. In October 2006 the Canadian Institutes of Health Research (CIHR) released a "Draft Policy on Access to CIHR-funded Research Outputs", which operates from the premise that "the primary purpose of all research in the public domain is the creation of new knowledge in an environment that embodies the principles of freedom of inquiry and unrestricted dissemination of research results" (http://www.cihr-irsc.gc.ca/e/32326.html). The draft policy calls for peer-reviewed articles based upon CIHR-funded research to be openly available online no longer than six months after publication. This is similar to the currently pending "Federal Research Public Access Act of 2006" in the United States (http://cornyn.senate.gov/doc_archive/05-02-2006_COE06461_xml.pdf). The US measure only applies to peer-reviewed journals. Critically, the Canadian measure *also* requires grant recipients to provide access to "research materials" and "research data". Research materials might include cell lines, nucleic acids, or research tools for evaluation. Research data would usually be an electronic data set, but could also include interview transcripts and survey results.

These research materials and research data are different examples of "grey content", which is the kernel of GL. If the Canadian policy goes into effect, and grey content is exchanged among researchers as a matter of course, this would revolutionize scholarly norms. The Canadian proposal is one example of the movement toward "open data" in science, which seeks to ensure maximum access to several types of data:

– *scientific data* "deemed to belong to the commons" (i.e., the human genome);
– *infrastructral data* such as that provided by geographic information systems;
– *factual data* that is not copyrightable (http://wwmm.ch.cam.ac.uk/blogs/murrayrust/?p=32).

[3] Conn VS, Valentine JC, Cooper HM, Rantz MJ. Grey Literature in meta-analyses. *Nursing Research* 2003;52(4):256-61.

[4] Benzies KM, Premji S, Hayden KA, Serrett K. State-of-the-evidence reviews: advantages and challenges of including grey literature. *Worldviews on Evidence-based Nursing* 2006;3(2):55-61.

[5] See ref. 3.

[6] ANSI/NISO. *Scientific and Technical Reports – Preparation, Presentation, and Preservation*. Bethesda, MD: NISO Press; 2005. (Standard Z39.18-2005). Available at: http://www.niso.org/standards/resources/Z39-18-2005.pdf; last visited Nov. 14, 2006.

GL producers and users should become strong proponents of the open data movement as well. The Open Data Foundation is one organization working in this area (http://www.opendatafoundation.org/). A leading proponent of the open data movement is Peter Murray-Rust of the University of Cambridge. Murray-Rust is also one of the developers of the concept of "datuments", which are scientific papers enriched with data elements that can only be accessed and interpreted electronically (http://jodi.tamu.edu/Articles/v05/i01/Murray-Rust/). Datuments integrate traditionally grey and white content into one package, and their proliferation would be strong evidence of the end of the distinction between grey and white literature. As argued at GL7,[7] this distinction would eventually fade with time. Time will tell whether all this is correct. In the meantime, paying careful attention to the "Nancy style" and ANSI/NISO standard is an excellent way to support the promotion of GL.

Reflections on the development of the "Nancy style"

The difficulties encountered during the creation of the *Guidelines* were analysed to reflect on the selection process required to reach consensus at international level, even within a small group of experts in a specific field. The major points of discussion mainly regarded the electronic GL and adoption strategy:

- **Electronic grey literature**
  The "Nancy style" is mostly paper oriented, because editorial consistency and ethical considerations recommended for traditional documents do apply also to digital publications. Yet, progressively more and more GL is being produced, stored, published and made available electronically and in order to manage relevant GL publications, metadata are required. The importance of metadata, as the natural evolution of library catalogue records, had been already stressed in the first version of the "Nancy style" (when dealing with report structure: Section 4.2 of the *Guidelines*), but no metadata schema was then provided since it was difficult to find a formula that would satisfy all requirements. At present, much GL is catalogued using the Dublin Core Metadata Standard (DC). However – as Keith Jeffery of the UK Council for the Central Laboratory of the Research Councils (CCLRC) pointed out working on the "Nancy style" draft – this standard suffers from several problems: a) it is machine-readable but not machine-understandable; b) it does not have a formalised syntax or semantics and therefore is open to ambiguous interpretations. Therefore, he proposed a formalised metadata standard (an umbrella standard, mainly generated from Dublin Core metadata: "Formalised DC" based on the concepts of the CERIF Model (www.eurocris.org/Cerif). Yet, as the traditional cataloguing practice has different rules, similarly different communities may adopt different metadata schema. Nowadays the World Wide Web provides the possibility to search for information across heterogeneous archives/databases/catalogues, but the systems managing different information resources must be "interoperable" (capable to work together), and interoperability requires that the same metadata schema be used. As Stefania Biagioni (of the Italian Istituto di Scienza e Tecnologie dell'Informazione - ISTI, Consiglio Nazionale delle Ricerche) clearly commented, there is much work towards standardization and the Dublin Core Initiative (http://dublincore.org/) is receiving worldwide consensus as it suggests adding a very simple metadata record to any specialized one.

- **Adoption strategy**
  When consensus was to be reached to release the first version of the *Guidelines*, a formal approval was asked to all organizations wishing to officially adopt them. Contrary to expectations, consensus was given only by a small number of institutions as the official adoption was sometimes a difficult step. Yet, support and encouragement did not lack: a less formal apprach in launching the *Guidelines* and getting them adopted was soon granted by all institutions involved in their creation. For example, a large international organization (Organisation for Economic Co-operation and Development - OECD), which took part in the

---

[7] Banks M. Towards a continuum of scholarship: the eventual collapse of the distinction between grey and non-grey literature. *Publishing Research Quarterly* 2006;22(1): 4-11.

development of the *Guidelines*, expressed concern to officially endorse them (and in fact, it did not), because that would require a great deal of internal debate and discussion with their own members. Suggestions were made to follow a voluntary system backed up by an official recognition of compliance to facilitate the adoption of the *Guidelines*. This would encourage like-minded supporters within an organisation to informally use the *Guidelines* and then gain the official "stamp of approval" to show that they are really following them. Actually, other organizations policies take a voluntary approach in the documents they recommend, such as the Association of Learned and Professional Society Publishers (ALPSP) with more than 230 not-for-profit publishers. As suggested by the OECD, voluntary sign-up is a less demanding step for organisations to take, but the effect is the same – more and more publishers will opt to use them.

In the next version of the "Nancy style" the adoption strategy may be revised, envisaging a voluntary approach followed by a subsequent official endorsement, once the organisation can show that *Guidelines* are actually followed.

Comparison between "Nancy style" and ANSI/NISO Z39.18

During the working process leading to the production of the *Guidelines*, and also in the period soon after their publication, we came in touch with other standards, guidelines, house styles, etc.

All useful documents were further studied in view of the future updating of the *Guidelines*. In particular, the *ANSI/NISO Standard Z39.18-2005 Scientific and Technical Reports – Preparation, Presentation, and Preservation* (released in 2005) has been considered a valuable source for comparison; therefore, we carefully analysed it and pointed out the main differences listing them under the following headings.

General considerations

The major differences concerning the two documents as a whole regard:

- **Document type**
  They are different in that the "Nancy style" represents *guidelines* – that is general principles agreed upon by a small group of experts, to be followed as an indication or outline of policy or conduct –, while the ANSI/NISO Z39.18 is a proper *standard*, developed by the Standards Committees of the US National Information Standards Organization (NISO), subject to rigorous control and approval process including peer review. This is why also the structure of the two documents is different since the standard may repeat concepts in different sections which may be used separately, while the *Guidelines* are intended as an easy to read document giving the general idea for recommended items. The *Guidelines*, different from standards, do not give full details on format and style.
  Moreover, the "Nancy style" represents *international* guidelines developed by a corporate author (GLISC), which worked on the draft proposed by the Istituto Superiore di Sanità, and signed approval of this best practice on behalf of their respective organizations, while the ANSI/NISO Z39.18 is a *national* standard approved by the American National Standards Institute through a number of Voting Members.

- **Paper *vs* digital document medium**
  The "Nancy style" is mostly paper oriented giving recommendations on report preparation mainly reflecting a traditional paper structure, while the organization pattern of the ANSI/NISO Z39.18 is user-based more than content-based. The key concepts incorporated in the American standard mainly refer to metadata, persistence of links, interoperability, creation, discovery/retrieval, presentation in digital format (DTD, XML, XSL), maintenance and preservation (original content, software and media); it also contains a metadata schema, which is absent in the *Guidelines*.

- **Annexes**
  All material included in the "Nancy style" is approved by the GLISC, while the ANSI/NISO Z39.18 provides a large amount of additional information (almost half of the pages) that is not

part of the Standard (Appendices including selected annotated bibliography, glossary, Dublin Core data elements, etc.).

Content considerations

In general, the "Nancy style" contains technical requirements for a report, but does not include full details (i.e. format, style, etc.); yet, it provides important elements, which are not present or not fully described in the ANSI/NISO Z39.18:

- **Ethical issues**
  An initial section is explicitly devoted to authorship, editorship, peer review, conflicts of interest, privacy and confidentiality.

- **Instructions to authors**
  Producers are strongly recommended to issue instructions to guide authors in the production of a formally correct document containing ethical and editorial issues as well as indications for formats, styles, illustrations, etc.

- **Revision**
  Special attention is given to revision editing as GL is not generally peer reviewed, or produced with editorial support; therefore, it is fundamental that authors be aware of the importance of a careful revision of their texts before diffusion.

- **Reference style**
  The adoption of the "Vancouver style" is recommended and examples and rules are given as a fundamental step for information retrieval.

As regards document structure, it is basically the same in "Nancy style" and ANSI/NISO Z39.18, with minor terminological variations. Yet, the American standard explicitly gives indication on:

- Report Documentation Page (since it is used by some agencies within the federal government, and also some sample pages are given).
- Distribution list.
- Glossary (although not part of the Standard).
- Executive abstract.

Technical recommendations

Since the "Nancy style" represents guidelines and not a standard, all technical considerations are limited to the essential, while the ANSI/NISO Z39.18 gives indications (all absent in the "Nancy style") on:

- **Print-specific/non-print-specific recommendations**
  The Section 6 "Presentation and display" describes standard methods for ensuring consistency in presentation including designing visual and tabular matter, formatting, etc. and makes a distinction between rules applicable to all reports regardless of mode of publication (paper or digital) and rules applicable to reports published in paper form only.

- **Format**
  Specific information is provided on fonts, line length, margins, page numbering, style, units and numbers, formulas and equations, paper (format and type), printing equipment, ink.

The ANSI/NISO Z39.18 also includes specifications on index entries and errata, which are not present in the "Nancy style".

Promotion of the "Nancy style"

When the "Nancy style" was released, we tried to promote it at best through all possible channels: announcements circulated in different forums; a logo for GLISC was created and its site was developed (now appearing high ranked in Google and also generating many links); translations in French (by the Institut de l'Information Scientifique et Technique) and Italian (by the Istituto Superiore di Sanità) were made available; and training courses were held using the *Guidelines* as a reference tool.

All these initiatives permitted to spread information about GL in environments where very little was known on this emerging communication channel or it was completely disregarded. Most often, people realized that they used and produced GL without being aware of all the implications laying behind it.

In particular, within the world of scientific editors, much interest in grey matters was shown by professional editors, who were generally unaware of all concerns relating to GL. Following such interest, they even asked to write a chapter on GL for the *Science Editors Handbook*.[8] This book is edited by the European Association of Science Editors (EASE), the major professional association of editors in Europe, defining itself as "an internationally oriented community of individuals from diverse backgrounds, linguistic traditions and professional experience who share an interest in science communication and editing". This chapter was most welcome by all members of the Association and they also asked to spread more information on GL through the journal *European Science Editing*.[9]

The possibility was also envisaged that the new born "Nancy style", like "Vancouver style", might develop into a *de facto* standard, representing uniform requirements for the production of technical reports produced by an international group (GLISC), in the effort to combat ignorance of the best editorial practices and permit integration, cooperation and standardization. Following such enthusiasm, a proposal was made that the direct promotion of the "Nancy style" by EASE might be included in the wider programme of penetration into the world of European science editing, even if GL is traditionally considered as a *sui generis* editorial product.

In the scientific arena, where career is guaranteed by publications in high impact journals, GL plays an important role not only as a vehicle of information, but also as a first step towards autonomous productions of editorially sound documents. Training inexperienced authors to allow a correct production of grey material may represent the best way to reach publications of higher levels. Tutoring in scientific writing and data presentation, in fact, also helps improving research methodology. The experience, gained in years of writing courses in the biomedical field, gives evidence that when scientists understand the reasons why editorial rules and standards must be applied (for both grey or white production), their publication output will rapidly improve as well as the challenges to have an article accepted by mainstream journals.

To date, most of the promotional efforts for the "Nancy style" have occurred in Europe. At the GL8 Conference in New Orleans, authors of this paper will discuss how to increase awareness of the "Nancy style" in North America.

Proposals for updating the "Nancy style"

From the analysis of the above described actions (Analysis of GL in the changing context; Reflections on the development of the "Nancy style"; Comparison between "Nancy style" and ANSI/NISO Z39.18; Promotion of the "Nancy style"), we propose the following suggestions to update the *Guidelines*, which we recognise to be the best tool to promote awareness and empowerment of GL creators:

- **Adding an Appendix on metadata**
  As regards electronic GL, the adding up of an Appendix containing metadata based on Dublin Core is proposed. The Appendix will give a very simple structure representing a general model to follow without going into details. An example for descriptive (bibliographic)

---

[8] De Castro P, Salinetti S. Writing and issuing Grey literature. Chapter 1-3.6. In: *Science Editors Handbook*. European Association of Science Editors; 2006.
[9] De Castro P. Scientists produce and use grey literature, but are they aware of the implications of doing so? *European Science Editing* 2006;32(4):95-7.

metadata regarding an entire issue and a contribution is given in Table 1. This possible form (based on Biagioni's suggestions) will be discussed at GL8. These metadata should be integrated with administrative (rights and software) and structural metadata (hierarchical levels, information used to display and navigate digital resources, etc.), according to the ANSI/NISO Z39.18 structure.

Table 1. Descriptive metadata schema suggested for an entire issue and a contribution
        (**to be discussed at GL8)**

Type of document	Descriptive metadata based on Dublin core
Entire issue	dc:title dc:date.issued dc:date.issued.human dc:relation.ispartof.publication.desc dc:relation.ispartof.publication dc:relation.ispartof.volume dc:relation.ispartof.year dc:publisher dc:contributor dc:source dc:format
Contribution in an issue	dc:title dc:creator dc:identifier dc:description.fulltext dc:description.abstract dc:format.extent.pagination dc:date.created dc:date.human dc:type dc:relation.ispartof.publication.desc dc:relation.ispartof.publication dc:source dc:relation.ispartof.volume dc:relation.ispartof.issue dc:relation.ispartof.year dc:subject dc:subject.msc dc:subject.jel dc:publisher dc:contributor dc:language dc:relation.ispartof.publication dc:format.extent.pagination

- **Creating a Subject index**
  A Subject index will be very useful to facilitate retrieval of specific items although the online availability permits an easy navigation through the text.

- **Providing more technical advice on digital format**
  Although the philosophy guiding the production of paper documents is unchanged in digital ones, the suggestions given by the ANSI/NISO Z39.18 in Section 6 "Presentation and display" represent a good example to follow as a distinction is made between rules applicable to all reports regardless of mode of publication (e.g., paper, CD-ROM, or Web) and rules applicable to reports published in paper form. A simple table will be added for easy readability including mandatory or optional elements with print-specific and non-print-specific indications.

- **Facilitating reference**
  Since GL is often incorrectly cited, it should be useful to give an indication on how to cite the document using the phrase "To be cited as" followed by the correct citation.

Final considerations

As promoters of the *Guidelines* and members of the GLISC, the experience of working with national and international realities permitted us to reflect more closely on the importance of creating a useful reference document applicable to different contexts.

The *Guidelines* are to be considered as a suggested model rather than a model in itself; they represent a basic step to improve quality in the different stages of GL production in view of its wider electronic circulation. The proposals for their updating will make them more effective, although a regular revision is required to keep pace with the changing ITC scenarios and information policies.

Traditionally editorial rules and ethical considerations were disregarded in the production of GL with negative implications on its quality. The members of the GL community should promote their diffusion mainly within GL issuing organizations that are less aware of existing standards regulating GL production and distribution.