Knowl. Org. 37(2010)No.3
C. Gnoli. Classification Transcends Library Business

223

Classification Transcends Library Business

Claudio Gnoli

Università di Pavia. Biblioteca della Scienza e della Tecnica.

Sezione di Matematica, via Ferrata 1, I-27100 Pavia, Italy <gnoli@aib.it>

Claudio Gnoli has been working as an academic librarian since 1994. His main interest is classification theory. He has published papers on this subject in several international journals and conference proceedings. He is member of the scientific advisory boards of the Universal Decimal Classification Consortium (UDCC) and of the journal *Knowledge Organization*, and vice-president of the International Society for Knowledge Organization (ISKO).

Abstract: Although bibliographic classifications usually adopt a perspective different from that of object classifications, the two have obvious relationships. These become especially relevant when users are looking for knowledge scattered in a wide variety of forms and media. This is an increasingly common situation, as library catalogues now coexist in the global digital environment with catalogues of archives, of museums, of commercial products, and many other information resources. In order to make the subject content of all these resources searchable, a broader conception of classification is needed, that can be applied to any knowledge item, rather than only bibliographic materials. To illustrate this we take an example of the research on bagpipes in Northern Italian folklore. For this kind of research, the most effective search strategy is a cross-media one, looking for many different knowledge sources such as published documents, police archives, painting details, museum specimens, organizations devoted to related subjects. To provide satisfying results for this kind of search, the traditional disciplinary approach to classification is not sufficient. Tools are needed in which knowledge items dealing with a phenomenon of interest can be retrieved independently from the other topics with which it is combined, the disciplinary context, and the medium where it occurs. This can be made possible if the basic units of classification are taken to be the phenomena treated, as recommended in the León Manifesto, rather than disciplines or other aspect features. The concept of bagpipes should be retrievable and browsable in any combination with other phenomena, disciplines, media etc. Examples are given of information sources that could be managed by this freely-faceted technique of classification.

1.0 Bibliographic classifications and objects classifications

Classification schemes developed in the context of library and information science are often referred to as bibliographic classifications. This label implies an important assumption: that the subjects indexed are considered only inasmuch as the content of some bibliographic material. In other words, the scheme does not classify objects as such, but objects referenced by documents. As documents can be about anything, this does not impose any limit to the scope of bibliographic classifications; but it does impose a particular perspective; that is, the perspective of being treated in documents, in combination with other subjects, in the light of some viewpoint. For this reason, bibliographic classifications are said to be aspect classifications (Svenonius 2000; Slavić 2007).

Most aspect classifications are based on disciplines: each subject is taken as treated under the perspective of a given disciplinary domain of knowledge. A book dealing with animals can be classified in zoology if it deals with animals in the way typical of naturalistic studies, or in animal husbandry, if it deals with them from the viewpoint of their production value, or in arts, if it shows artistic illustrations of animals, and so on. The classifier is thus forced to decide under which discipline to subsume the document about animals; and does not have the option to index the document's content as being about "animals." While disciplines are a traditional, widespread way of organizing knowledge, in many cases they also act as a superstructure adding an unnecessary perspective to the content itself. Indeed, knowledge increasingly develops in interdisciplinary ways: a subject studied using methods from several concur-

224

Knowl. Org. 37(2010)No.3
C. Gnoli. Classification Transcends Library Business

rent disciplines, or by a newly developed discipline not yet provided for in the scheme, can thus be represented in inaccurate and partial ways, resulting in a hindrance for cross-disciplinary research (Austin 1969; Beghtol 1998; Szostak 2008).

In contrast to this, classifications of objects are aimed at directly organizing objects, irrespective of their treatment in any document. The organization of animals in a zoo, or in the fossil rooms of a natural history museum, or in the catalogue of a pet shop, will be classified simply into groups of animals, rather than into branches of zoology or some other discipline. Disciplines have no relevance here, as what is organized are the objects themselves, not the way in which they may be treated in documents. Classifications of objects are often called taxonomies: the scheme used to arrange literature on knowledge organization (Dahlberg 1993) lists "objects classification systems (taxonomies)" in a class separate from "[bibliographic] classification systems and thesauri."

Object classifications have long been used in different ways, forming a very important systematic part of many sciences, like astronomy, medicine, or archaeology. Recently, relatively simple taxonomies of objects or processes have also been used in business information systems and in website menus, becoming popular among knowledge managers and information architects (Lambe 2007). A familiar application are the web catalogues of vendors, allowing their users to select products, like clothes or cars or wine bottles, by searching for them using several "facets" and browsing the resulting lists. Although not always presented formally as such, these applications are based on classifications of objects.

Although they belong to different traditions of application, bibliographic classifications and object classifications nevertheless have obvious elements in common. After all, both are aimed at organizing knowledge items, irrespective of whether these are material objects or their textual or pictorial representations. Aspect does matter as a component of classification; still it is an additional dimension, superimposed on objects, that form the original basis of any knowledge. Indeed, bibliographic and scientific classifications are sometimes considered together as a topic (Kyle 1962). In order to list the objects of a particular discipline, bibliographic classifications obviously draw on scientific classifications, although defining their main hierarchies, auxiliary hierarchies, and facets in a disciplinary perspective. In turn, bibliographic classifications and thesauri, despite their

aspect bias, are used sometimes to organize artifacts (MDA 1992). UDC also "is currently used for subject indexing in a major private art collection, because paintings, sculptures and ornaments can yield historical information about persons, costume, animal breeds, vanished buildings or landscapes, and many other subjects" (UDC 1999).

I suggest that such relationships between the bibliographical and the objects should not be viewed as accidental, but as important clues on the general nature of classification. Understanding these relationships in the light of a broader view of classification is becoming even more important, as we often browse across knowledge sources in different forms and carriers. Indeed, related information is increasingly available in more than one technological form (cross-mediality), e.g. a film, a book, a website, and various gadgets all on the same story. What is most relevant is not the carrier, which can easily evolve or be crossed to another medium: it is the content. The opportunity to integrate knowledge from different sources calls for connections between different information systems in order to make related information retrievable by a single search or through a few links. As the formal and aspect dimensions vary, common patterns can only be found in the semantic dimension, that is, in the phenomena about which the different sources convey knowledge.

In particular, the digital media are now pushing institutions towards an integration and standardization of records coming from libraries, archives and museums (Rayward 1998). These are indeed among the major sources of knowledge, be it "bibliographic" or "archival" or "museological." More general models thus appear to be needed for the management of knowledge items. Extended notions of "work" and its "instantiations" as knowledge entities, encompassing not only published documents, but also archival pieces or archaeological artifacts, are explored by Smiraglia (2001; 2007). The CIDOC Conceptual Reference Model is proposed to "provide the "semantic glue" needed to mediate between different sources of cultural heritage information, such as that published by museums, libraries and archives" (CIDOC 2006). Recently, this trend is also followed by the International Federation of Library Associations (IFLA) in developing the conceptual models of the Functional Requirements for Bibliographic Records (FRBR) family. These increasingly refer to knowledge items treated by archives and museums, beside libraries (Tillett 2008). IFLA has begun to collaborate with the professional associations devoted to paper archives, to audiovisual

Knowl. Org. 37(2010)No.3
C. Gnoli. Classification Transcends Library Business

225

archives, and to museums. They all feel concerned with what is now called the documentary heritage, to promote a unified approach to the management of its records in memory institutions. For example, the recently constituted Library and Archives Canada (Parent in prep.) presents itself as "Canada's national collection of books, historical documents, government records, photos, films, maps, music... and more."

2.0 The case of northern Italian bagpipes

In order to explore the connections between different forms of knowledge, and the ways in which they can be managed by classification, we will consider a specific real case, of which the author has direct experience. It is an investigation about the forms, diffusion, and uses of bagpipes in the folk cultures of northern Italy. This appears to be a very concrete and potentially rich scientific subject. At the same time, however, it would be very difficult to study it by the standard tools of academic libraries, including catalogues using a disciplinary classification. Indeed, very few documents dealing with this specific subject are published within the literature of a single discipline (musicology probably being the most obvious candidate).

One reason for this situation is that the subject is poorly known in literature, as much of folk culture in northern Italy has been erased by the modern lifestyle that has been spreading during the last century. Bagpipes and their music thus have completely disappeared from the whole Po Plain and from most of the Alps. Only a few mountainous regions in the central Alps and the northern Apennine still keep some local bagpipe varieties (baghèt, musa, piva), which have been built again recently for use in folk revival. Therefore, searching for information on northern Italian bagpipes in practice also means doing new research on them, at least by assembling information currently scattered in many places. Such creation of new knowledge by identifying meaningful connections between available sources, otherwise unrelated, is indeed a major application of knowledge organization systems (Davies 1989; Beghtol 1995; Szostak 2008).

Making this search possible in significant collections, however, would require that the different sources were indexed by some common descriptor, which is hardly the case at present. Sources of knowledge on folk music are typically scattered in a number of different media, often not indexed by their subject content. Ethnomusicological archives including audio

and video recordings are often uncatalogued, or their catalogues are not available online. The task of indexing them also questions some common assumptions of classification, as music traditions would need to be described as for their distinctive features rather than their similarities, while the common classifications by genre do not apply across different cultures (Proutskova 2007). This is even more true for non-bibliographic materials like police reports or frescos, that are not managed primarily as sources of musical knowledge, but once found can prove very useful as such.

Besides the few available books and papers dealing with Northern Italian bagpipes, the following sources proved useful to this research:

a) reference to bagpipes in traditional tales. These would be indexed in a disciplinary classification under philology or ethnography, or can be unpublished and have to be found directly in field notes and recordings of interviews with people native of the relevant regions;
b) description of customs. Bagpipes were often played on holy days, country feasts or country fairs, so that accounts of these can prove their presence at a given time and place. Such kind of descriptions can be found in old books describing a region and its inhabitants from many aspects, where music players could be mentioned only occasionally, rather than be the main subject;
d) travel reports written in the past in foreign languages by authors coming from urban milieus and from more learned countries (France, Germany) to visit the regions of interest;
e) police archives of the relevant administrations, mentioning players or dancers who are involved in arrests, e.g. following a brawl during a feast. In the little rural municipalities of the relevant region, such archives are often difficult to access, and usually lack any kind of index;
f) presence of bagpipes in painting and illustration details. These are often the only documents testifying the presence of bagpipes before the 19th century. Details are generally faithful, allowing the reconstruction of some organological features of the instruments. For example, frescos dating back several centuries painted in three churches in the Piedmont Alps have proved the presence of bagpipes in an area where no other proof of it exists; details of one of these frescos

226

Knowl. Org. 37(2010)No.3
C. Gnoli. Classification Transcends Library Business

have been used to build a real working instrument with the same features, that can now be played and studied. Bagpipes could be represented in other existing pictures as well, but finding them would require that the objects represented in them be indexed somewhere, however prominent they are in the whole picture. This is not a common practice in picture catalogues, although some systems are available like Iconclass, that has main classes for "nature," "human being," "society," "abstract ideas," and "history," including events and persons;

g) bagpipe specimens kept in museums across the world. Ethnomusicologists have collected and studied bagpipes that are now kept in museums of popular cultures or of musical instruments, often far away from the place of their origin. One of the very last bagpipes (musa) having played in the Piedmont Apennine is now kept in the National Museum of Scotland in Edinburgh. Users looking for information on them would gain much benefit by finding a record of it, but museum catalogues are not interconnected with library catalogues. Furthermore, local bagpipe types often lack standard names, being referred to by words in different local dialects, and information on the details and the geographical origin of the specimens is often not searchable;

h) other objects kept in various collections and museums, like a puppet of a bagpipe player by the historical Lupi puppet acting company in Turin, Piedmont (again, a region where no bagpipe tradition is currently known);

i) organizations dealing with folk dances. Local organizations active in folk revival of traditional dances often hold knowledge about old instruments playing such music, that is unpublished or published in items with a very limited distribution, and are not found in most library or bookshop catalogues. Although organizations are not "documents" in the usual sense, they are knowledge sources on specific topics, hence potentially indexable by a classification scheme. Indeed, the Broad System of Ordering, developed with the involvement of bibliographic classification experts, was aimed at classifying institutional sources in a world-wide information network, rather than documents in the standard sense (Coates et al. 1978).

3.0 Phenomena as knowledge unit

Many of the difficulties reported above may be further complicated by a narrow conception of classification (or, more generally, of knowledge organization systems). For instance, if we would conform strictly to the definitions of some bibliographic classifications, only the subject content of published, non-fictional documents would be indexed, and these would be partitioned into a fixed number of disciplinary fields. Such an approach may rule out important knowledge sources like archival documents, museum specimens, artistic expressions, organizations devoted to the subject, etc. Although such narrow definition has good historical reasons and descends from the solid tradition of library classification, it is our thesis that we would benefit from a broadened conception of classification, able to encompass a greater set of knowledge items, like those described above. These all seem to be potentially covered by the field of knowledge organization (Gnoli 2008, question 1), as they are taken as sources of knowledge, rather than just as everyday objects.

If we are to focus on the knowledge content shared by the different sources, clearly we should not base our system on their formal or aspectual features, like the material details of the carrier or the perspective adopted. Rather, what is common to cognate knowledge items are primarily the objects documented by them. What books on bagpipes, pictures with bagpipes, recordings of bagpipes, and organizations dealing with bagpipes have in common are bagpipes—not musicology or police history, not bibliographic or museographic metadata. In other words, the needs of intermediality push us to base our classifications on the phenomena that are recorded and described, in whatever perspective and form their knowledge is expressed and kept.

Phenomena (the object of knowledge) are one of the basic dimensions of knowledge, together with aspect (the perspective by which knowledge is treated and developed) and carrier (the material medium in which knowledge is transmitted). All these three dimensions can be represented in a classification, as each of them can be relevant for some search. However, the León Manifesto recommends that they be represented separately, and that priority be given to phenomena (ISKO Italia 2007; Szostak 2008). Indeed, phenomena seem to be the most suitable reference to be taken as a knowledge unit, as the other dimensions can be defined and represented in

Knowl. Org. 37(2010)No.3
C. Gnoli. Classification Transcends Library Business

227

terms of them: a discipline is a cultural phenomenon, a carrier is an artifact phenomenon aimed at conveying knowledge.

This is the approach taken in the Integrative Level Classification (ILC) project, currently experimenting with a knowledge organization system based on phenomena. Experience using a version of UDC interpreted as referring to phenomena, rather than disciplines, is also reported (Cousson 2009). Phenomenon classifications are better suited for application to museum specimens, to organizations, to art works, as well as to books and papers. In particular the notion of discipline, while making sense when applied to most academic books and papers, does not do so when applied to the subject of a police report, a tale, or a fresco (note that classifying tales under "literature" and frescos under "arts" would index them in terms of their carrier, not their subject content).

4.0 Freely faceted classification

The recommendations of the León Manifesto look consistent with the purposes of a broadened conception of classification as outlined here. Indeed, they would make possible the retrieval of each phenomenon, or each aspect, or each carrier independently of the other facets and dimensions with which it is combined, all represented by separable notations. We are thus able to index a knowledge item as "bagpipes: Northern Italy: museums", and another as "bagpipes: Northern Italy: books", or as "players: bagpipes: Piedmont: frescos", or as "feasts: bagpipes: Genua province: 19th century: archives". All these subject combinations would be retrieved in a search for "bagpipes."

This kind of free combination of phenomena in classification is made possible by a freely faceted structure, of the kind first described by the Classification Research Group, especially in the writings of Derek Austin (Hopkins 1973; Austin 1976), and currently developed in the ILC project. Indeed, in freely faceted classification, the citation order of combined concepts (facets) does not depend on the context of a given discipline anymore. The classmark representing a knowledge item starts by the notation representing the phenomenon that is its main theme, and the other phenomena are expressed in terms of their relationships to the first one.

"Theme" is a linguistic notion referring to the part of a sentence, usually expressed first in most languages, that sets the context for more information given about it in the rest of the sentence, like "feasts"

in "As for feasts, these were in the country, with dancers and bagpipe players." Within the IFLA working group on Functional Requirements for Subject Authority Record, "thema" is proposed as a fundamental syntactic entity, in a model of verbal subject indexing also inspired to facet analysis (Buizza et al. 2009).

For example, a picture representing a country feast which includes dancers and bapgpipe players could be classified primarily under the main theme "feasts". The other elements, like "country," "dancers," "players," "bagpipes" will be connected to it as its facets. Country is a place facet of feasts; dancers and players are agent facets of feasts; and bagpipes are an instrument facet of players. The resulting compound class is not bound to any particular discipline: its place in a systematic display will depend only on the phenomenon taken as the lead, in this is case feasts, that will be listed with socio-cultural phenomena. Nevertheless it can be combined with phenomena belonging to other integrative levels of reality, like country (level of ecosystems), dance and music (level of art works), and bagpipes (level of artifacts). These could in turn take the lead position if they happened to be the main theme of another knowledge item, like in the case of a specimen kept in a museum of musical intruments, that would be represented primarily as "bagpipes," maybe with "dance" as a function facet, and "country" as a provenance facet.

The phenomenon "bagpipes," expressed by a constant notation, will thus be combined with other phenomena according to their relationships in each occurrence in any kind of knowledge item. In a browsable list, knowledge items will be listed primarily under their lead phenomenon, so that only those where bagpipes are the main theme will be listed under "bagpipes." However, the user looking for information on bagpipes will be able to extract it also from any combination where a "bagpipes" facet appears.

The limitations of some disciplinary classifications, deriving from their assumption that knowledge items are only bibliographic materials expressing a disciplinary perspective, can thus be overcome because the different contexts where a concept occurs are not lost, as they are expressed by their relationships with other phenomena, with aspect, and with carrier.

5.0 Conclusion

The tradition of library and information science undoubtedly represents the richest heritage of principles and techniques of knowledge classification. However,

228

Knowl. Org. 37(2010)No.3
C. Gnoli. Classification Transcends Library Business

its origins, which are to a large extent aimed at organizing printed academic documents, mean that most classification systems have been designed with bibliographic and disciplinary assumptions built in, which are now impossible to separate from the representation of phenomena that they contain.

As we have seen, adaptations of bibliographic classification to non-library environments, or to non-disciplinary treatments, have also been performed; the result are hybrid systems that would probably reveal themselves problematic if analyzed formally. In any way, these applications indicate a need for an extended notion of classification, that can be applied not only to bibliographic materials, but to any knowledge item, that is, anything kept and shared in order to convey knowledge. Not only a printed document can convey knowledge, but also a multimedia resource, a find in a museum, or a living plant with its label in a botanical garden. What is really important about them are not their formal aspects, but their knowledge content. This should be given priority in classification, if we are to make possible a better integration of the huge diversity of knowledge sources available today. The task of classification, and of subject indexing in general, is to normalize the subject content of knowledge items, making them searchable independently from the form and context in which they occur.

This requires that we move on from the traditional view of classification, bibliographic and disciplinary, to a generalized view, that is not a negation of the former, rather an evolution of it. The new media, and the need for integration of different forms of knowledge, call for a transition from classification as developed in libraries, where any item has to be forced within some disciplinary hierarchy, towards knowledge classification, taking phenomena as its main units and exploiting the flexibility of freely-faceted structures, that can be applied to a wider range of knowledge forms and situations.

References

Austin, Derek. 1969. Prospects for a new general classification. *Journal of librarianship* 1(3): 149-69.

Austin, Derek. 1976. The CRG research into a freely faceted scheme. In Arthur Maltby ed., *Classification in the 1970s: a second look*. London: Bingley, pp. 158-94.

Beghtol, Clare. 1995. 'Facets' as interdisciplinary undiscovered public knowledge: S.R. Ranganathan in India and L. Guttman in Israel. *Journal of documentation* 51: 194-224.

Beghtol, Clare. 1998. Knowledge domains: multidisciplinarity and bibliographic classification systems. *Knowledge organization* 25: 1-12.

Buizza, Giuseppe, Bultrini, Leda and Cheti, Alberto. 2009. Subject analysis and indexing: an "Italian version" of the analytico-synthetic model. *Proceedings IFLA satellite preconference Looking at the past and preparing for the future, Florence, 20-21 August 2009*.

CIDOC, upd. 2006. *The CIDOC Conceptual Reference Model.* Foundation for Research and Technology, Hellas. Available http://cidoc.ics.forth.gr.

Coates, Eric, Lloyd, Geoffrey and Simandl, Dusan. 1979. *The BSO manual: the development, rationale and use of the Broad System of Ordering*. The Hague: FID. Excerpts republ. at http://www.ucl.ac.uk/fatks/bso/manual.htm: University College London, 2002.

Cousson, Philippe. 2009. UDC as a non-disciplinary classification system for a high-school library. *Extensions & corrections to the UDC* 31 [in preparation].

Dahlberg, Ingetraut. 1993. Classification system for knowledge organization literaure. *Knowledge organization* 20: 215-22.

Davies, Roy. 1989. The creation of new knowledge by information retrieval and classification. *Journal of documentation* 45: 273-301.

Gnoli, Claudio. 2008. Ten long-term research questions in knowledge organization. *Knowledge organization* 35: 137-49.

Hopkins, Fran. 1973. General classification theory: the work of the Classification Research Group. *Library resources and technical services* 17: 201-10.

ISKO Italia. 2007. *The León manifesto*. Republ. *Knowledge organization* 34: 6-8. Available http://www.iskoi.org/ilc/leon.htm.

Kyle, Barbara R. F. 1962. Classification: an interdisciplinary problem: proceedings of an Aslib conference, London, 6th April 1962, *Aslib proceedings* 14: 222-62.

Lambe, Patrick, 2007, *Organizing knowledge: taxonomies, knowledge and organization effectiveness*. Oxford: Chandos.

MDA: The Museum Documentation Association. Terminology Working Group. 1992. *Thesauri for museum documentation: the proceedings of a workshop held at the Science Museum, London, 24 February 1992, Cambridge: MDA*.

Parent, Ingrid. 2010 [in prep.]. Convergence and integration among libraries, archives and museums: an international trend. In *Proceedings of conference*

Knowl. Org. 37(2010)No.3
C. Gnoli. Classification Transcends Library Business

229

Il mondo in biblioteca, la biblioteca nel mondo, Milano, 12-13 marzo 2009. Editrice Bibliografica, Milano.

Proutskova, Polina. 2007. Musical memory of the world: data infrastructure in ethnomusicological archives. In *Proceedings 8th International Conference on Music Information Retrieval, September 23rd-27th 2007, Vienna*, ISMIR. Available http://ismir2007.ismir.net/schedule.html.

Rayward, W. Boyd. 1998. Electronic information and the functional integration of libraries, museums and archives. In Higgs, Edward ed. *History and electronic artefacts*. Oxford University Press, pp. 207-24.

Slavić, Aida. 2007. On the nature and typology of documentary classifications and their use in a networked environment. *El profesional de la información* 16: 580-89. Also in *DLIST*, http://dlist.sir.arizona.edu/2097/.

Smiraglia, Richard P. 2001. *The nature of "a work": implications for the organization of knowledge*. Lanham: Scarecrow.

Smiraglia, Richard P. 2007. When is a terracotta hut urn like a sailor's deck-log?: meaning instantiated across virtual boundaries. In *Proceedings Museums and the Web 2007, San Francisco*. Available http://www.archimuse.com/mw2007/papers/smiraglia/smiraglia.html.

Svenonius, Elaine. 2000. The intellectual foundation of information organization. Cambridge (MA); London: MIT Press.

Szostak, Rick. 2008. Classification, interdisciplinarity, and the study of science. *Journal of documentation* 64: 319-32.

Tillett, Barbara. 2008. *The bibliographic universe and the new IFLA cataloguing principles: lectio magistralis, Firenze, 14 March 2008*. Firenze: Casalini libri.

UDC: Universal Decimal Classification, 1999, *pocket edition*, London: BSI.

All URLs are last checked in February 2010.