

BIC Standard Subject Categories – an Overview

November 2010

History

In 1993, Book Industry Communication (BIC) commissioned research into the subject classification systems currently in use in the book trade, with a view to establishing an industry-wide standard for assigning the broad subject category of books, especially for use with electronic data exchange applications.

BIC established a working party including representatives of publishers, booksellers, libraries and bibliographic agencies, and work began on formulating the BIC Standard Subject Categories, based on the proprietary classification scheme developed by Book Data. Howard Willows of Book Data (now Nielsen BookData) was appointed Chair and has remained so ever since.

The Final Draft version was released in March 1997. There followed:

- Version 1 in July 1998
- Version 1.1 in August 2002 (interim release)
- Version 2 in April 2006

A further interim release, Version 2.1, comprising only additional categories and modifications to existing headings (no deletions or relocations), is scheduled for release in late 2010.

Scope and Purpose

The initial purpose of the BIC Subject Categories was to provide:

- a hierarchical codified list of subject-based headings,
- for use in the UK book trade,
- which reflected the full range of publishing. including specialist areas,
- covering traditional subjects as well as the latest developments,
- providing a level of detail to each area that appeared neither impenetrable to the layman nor superficial to the expert,
- and which was accessible and simple to use without extensive training or classification experience.

Subsequent revisions sought to reduce the UK bias where possible to position the scheme as a global standard for the English language book trade. It has subsequently been adopted, either in its original form or as the basis of a local variant, in several countries, including Australia, Sweden, Spain, Italy, Greece and Russia. It is currently being considered for use in a project to develop a pan-Arabic bibliographic database.

The usage of the BIC scheme was intended as the classification/categorisation (see below) of books within the information supply chain from the publisher to the retailer, typically through a bibliographic aggregator, but it was not necessarily expected to be used by or visible to the consumer. Rather it was assumed that, in presenting information and products to customers, retailers would manipulate or map the BIC values to some extent, either to generate simpler shelf labels or to drive and underpin

more intuitive search terms. Thus different retailers are able to present products to consumers in their own way, avoiding any institutionalisation of the book retail trade.

The BIC scheme was not expected to be adopted by the library sector. It has, however, been used as the basis for the e4libraries subject headings scheme designed for use in libraries¹.

The BIC scheme was not specifically designed for sales analysis, but again has been used as the basis for Nielsen BookScan's Product Class scheme, used precisely for this purpose.

The BIC scheme was not designed particularly with regard to compatibility with the BISAC scheme, prevalent in North America, or with future convergence between the two schemes in mind. Nonetheless, developments in BISAC have always been examined as part of the BIC review process and the schemes are sufficiently similar to support reasonably accurate mappings or crosswalks.

The BIC scheme has thus been described as a kind of *lingua franca* for trading partners exchanging information about books, from and to which other schemes may be converted.

The benefits of an industry-standard set of subject categories are numerous. They will facilitate consistent communication of rich metadata between all sectors, support analyses of publishing output by subject, and enable meaningful comparisons of data from various sources within the trade.

Structure

The basic structure and principles of code construction have remained unchanged over the revisions. The scheme comprises some 2600 subject Categories arranged in 18 sections defining broad subject areas, plus a separate Children's section that covers all subjects; and over 900 Qualifiers covering 5 distinct additional aspects. All codes are hierarchically arranged, and of various length depending on their position in the hierarchy.

The codes for Categories are from 1 to 7 characters (4 alpha, 3 numeric)

The codes for Qualifiers are from 2 to 6 characters (1 numeric, 6 alpha)

Each of the Subject Categories sections is indicated by a single letter code, known as a Level 1 code, e.g.:

| | |
|----------|------------------------------------|
| <i>F</i> | <i>Fiction & Related Items</i> |
| <i>M</i> | <i>Medicine</i> |

These sections are subdivided into various Level 2 Headings, defining major topics within a given subject area, e.g.:

| | | |
|-----------|---------------------|--|
| <i>AM</i> | <i>Architecture</i> | (i.e. as a topic within <i>A = The Arts</i>) |
| <i>JN</i> | <i>Education</i> | (i.e. as a topic within <i>J = Society & Social Sciences</i>) |

¹ <http://www.bic.org.uk/51/E4libraries-Subject-Category-Headings/>

The amount of further subdivision varies from subject to subject, but the same principle applies whereby an extra character is added to the code for each heading that is a subdivision of a preceding heading. This ensures that the headings are arranged in a strictly hierarchical structure. Up to seven characters are permissible in a Category code, thus allowing for some fairly precise and narrow headings, although in practice only six characters have been used so far, e.g.:

HRCC97 *Quakers (Religious Society of Friends)*

Working back through this code character by character, HRCC97 locates this topic within *Protestantism & Protestant Churches*, within *Christian Churches & Denominations*, within *Christianity*, within *Religion & Beliefs*, within the broad section H for *Humanities*.

Alongside the subject headings are five sets of Qualifiers, which are used to indicate the appropriate Geographical area, Language and/or Time Period covered by a title, and also any Educational Purpose and/or Interest Age & Special Interest. These are also arranged hierarchically, and can be readily distinguished from the Subject Category codes by the fact that they begin with a number rather than a letter, e.g.:

1DV *Eastern Europe*
2LA *Australian Aboriginal languages*
3JJG *Inter-war period 1918-1939*
4KH *Designed / suitable for National Curriculum*
5AR *For reluctant readers (children)*

Application

The BIC Subject Categories will be applied to titles for a range of purposes, by book trade personnel working within various environments and with a range of computer systems. Some may permit more categories and qualifiers to be applied than others; some systems may impose limits on the number of characters allowed in a single code. The hierarchical structure of the codes can accommodate these variations, so long as some general rules are consistently followed:

- Ensure the first category entered is the primary subject (i.e. the most pertinent and important subject of the title);
- Assign as many categories as are relevant (space permitting);
- Classify titles as precisely as possible (i.e. using the highest Level category);
- Observe any usage notes and instructions linked to categories;
- Add Qualifiers whenever appropriate.

Advantages of the BIC scheme

The key features of the BIC scheme which have led to its widespread adoption and make it particularly suitable for online discovery of books and digital content are as follows:

Hierarchical coding structure

This enables searchers either to navigate a classification tree in a logical progression, opening up sub-categories as they go until they reach the most precise

topic for their needs. It also facilitates narrow or wild-card searches, and can be used to prompt searchers to move up or down the hierarchy.

Combinations of multiple codes

Rather than have ready-made categories that each incorporate a lot of characteristics of a topic, BIC supports a “building block” approach which is hugely flexible whereby any combination of Categories and/or Qualifiers can be used to specify precise topics.

Separate Qualifiers

In addition to their key role in combining with Categories to define countless topics, the Qualifiers can be used alone in searching to find, e.g. all books on China, whether on politics, history, art & culture, social aspects, food or travel guides.

Not strictly ‘Subject’ based

Although clearly and fundamentally a means of defining the “subject” of a book, the BIC scheme also allows for an appreciation of the intended usage and treatment of titles ostensibly on the same subject. Thus, in BIC a distinction is made between the scientific approach to the study of birds (*PSVW6 Birds (ornithology)*) and books for the enthusiast (*WNCB Birds & birdwatching*); between the social science of criminology (*JKV Crime & criminology*) and sensationalist accounts (*BTC True crime*), and so on.

Manageable size and level of detail

Although there is arguably some inconsistency as to the number of categories within subject areas across the scheme, the overall number of codes has remained fairly steady and has proved fit for purpose with regard to the needs of users and the ability to map to and from other schemes (DDC, BISAC, proprietary schemes).

Limitations of the BIC scheme

Qualifiers not assigned to specific subjects

It is a recognised limitation of the scheme that because there is no explicit link between a Category and a Qualifier – each Qualifier must be regarded as applying to the book as a whole rather than to any specific subject Category – there is room for ambiguity in the case of titles such as “Italian sculpture in French gardens”, which would carry exactly the same BIC values as a book called “French sculpture in Italian gardens”. Anyone searching for “Gardens + Italy” would retrieve both titles.

In most cases though – for instance, a travel guide to Madrid classed as WTH and 1DSE – this is not a problem as there is an implicit and obvious association even though no explicit link between the two values is made.

UK bias

Despite deliberate attempts to reduce the UK bias, and the global scope the scheme has always encompassed, there are still conspicuous examples where topics of particular UK interest are covered in more detail than others – most notably in the Geographical and Educational Purpose Qualifiers. But, this is in response to the needs of users, and the scheme can be expanded to give more balanced treatments as required, or at the request of users around the world.

Difficulty of Fiction

All bibliographic classification schemes struggle with sufficient granularity of Fiction categories, given the huge number of fiction titles published. There are typically

many, many more items classified under any Fiction category than under most other non-fiction headings. This problem is particularly acute with regard to novels that do not fit neatly into established genres.

Room for expansion within coding structure

The coding structure outlined above, although it contains the potential for many new values, is not infinitely extensible, and has already “run out of space” in certain subject areas, particularly at Level 5 of the Categories where numbers are introduced. Thus if more than 9 sub-divisions of a given Level 4 Category are required, they cannot be accommodated within the coding structure. Inevitably, those broad sections devoted to wide ranges of subjects are more prone to the exhaustion of possible codes.