LIBRARY OF CONGRESS SUBJECT HEADINGS
AND THE ASFIS THESAURUS

James W. Markham
Library - Cataloging Department
University of California
Santa Barbara, California 93106

Abstract

Much of the recent literature of aquatic sciences has been indexed using descriptors from the Aquatic Sciences and Fisheries Thesaurus (ASFIS Thesaurus). Items from this same literature, if cataloged by a library in the United States, normally have Library of Congress Subject Headings (LCSH) assigned to them. A comparison of the ASFIS Thesaurus and LCSH indicates that less than half the terms are common to both schemes. The comparison is discussed: the changing roles of descriptors and subject headings and the decreasing differences between them in the online computer environment.

INTRODUCTION

The literature of aquatic sciences and fisheries of recent years in journal articles, reports, and monographs has been indexed extensively by several agencies, but specifically by the Aquatic Sciences and Fisheries Information System (ASFIS), which focuses on this particular literature. Indexing by ASFIS employs descriptors from the ASFIS Thesaurus (1) and is done to facilitate retrieval in the Aquatic Sciences & Fisheries Abstracts database (2). Much of the same material is also cataloged in libraries. Although most journal articles will usually not be cataloged, most of the monographs and reports, many of them dealing with topics as narrow and specific as the journal articles, are cataloged. If cataloged by a library in the United States, these items will normally have Library of Congress Subject Headings (LCSH) (3) assigned to them.

Since much literature thus has two sets of terms assigned to it, by two different groups, it is of interest to determine how these terms compare. This study compares the ASFIS Thesaurus with LC Subject Headings.
METHODS

The ASFIS Thesaurus classifies descriptors in 12 separate facets (Table 1, Column 1) and these separate facet classifications were used for term matches. Using these separate lists rather than the main alphabetic ASFIS Thesaurus gives a total of approximately 6285 descriptors, assuming no descriptor appears in more than one facet. The largest facet is "Phenomena", representing 22.9% of all descriptors, followed by "Materials", 18.9%. The smallest facets are "Constants", 0.6%, and "Residual Concepts", 0.4%. All descriptors in each of the 12 facets of ASFIS Thesaurus terms were matched against LCSH. The reciprocal - matching all LC Subject Headings relevant to aquatic sciences against the ASFIS Thesaurus - was not performed. Matches were recorded for each descriptor as one of six possibilities: 1. + = same term in both lists. In cases where the only difference was singular vs. plural or American vs. British spelling, the terms were recorded as +. 2. - = no equivalent term or concept found in LCSH. 3. ST = similar term in both lists, similar enough to be readily found or recognized. 4. DT = different term in the two lists, i.e., concept exists in both, but terminology is different enough to be difficult or impossible to find without a cross-reference guide or a wide knowledge of multiple vocabulary possibilities for a particular concept. 5. NT = ASFIS descriptor equivalent used in LCSH is a narrower term. 6. BT = ASFIS descriptor equivalent used in LCSH is a broader term. Results were translated into percentage of each type of match within each ASFIS facet.

RESULTS

All facets together showed a match rate ("+") of 31.15%, while 54.6% had no LCSH equivalent ("-"") (Table 1). The highest match rate was in the "Sciences and Technology" facet (56.09%), the lowest in "Residual Concepts" (12%). The highest rate of similar terms (ST), different terms (DT), and broader terms (BT) was in the "Operations" facet. Overall, 3.8% of ASFIS descriptors had narrower LC equivalents (NT), while only 0.4% had broader LC equivalents (BT).

Examples of "+" matches include terms such as EXPLORATORY FISHING, SCALLOP FISHERIES, CIRCADIAN RHYTHMS, SEAFLOOR SPREADING (= Sea-floor spreading), BEAUFORT SCALE, ARTIFICIAL REEFS, and COCCOLITHS. Examples of "-" matches, or no apparent LC Subject Heading equivalent, include BATHYMETRY, TSUNAMI PREDICTION, TIDAL AMPLITUDE, BIOACCUMULATION, WAVE PROPAGATION (and most other wave descriptors), SCHOOLING BEHAVIOUR, CARRAGEENINS, SECHI DISCS, ICE-WATER INTERFACE (and most other interfaces), and LIMITING FACTORS.

Of all ASFIS descriptors, 7.5% have similar LC equivalents (ST), with the highest number in "Operations" (16.4%) and the lowest in "Constants" (2.8%). (In the following
TABLE 1. ASFIS Thesaurus descriptors compared against LCSH. + = Term matches; - = No equivalent in LCSH; ST = Similar term in LCSH; DT = Different term is equivalent in LCSH; NT = Narrower term is equivalent in LCSH; BT = Broader term is equivalent in LCSH. All figures except #terms in %.

<table>
<thead>
<tr>
<th>Descriptor Facet</th>
<th>#terms</th>
<th>% total</th>
<th>+</th>
<th>-</th>
<th>ST</th>
<th>DT</th>
<th>NT</th>
<th>BT</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPERATIONS</td>
<td>751</td>
<td>11.95</td>
<td>23.43</td>
<td>43.54</td>
<td>16.38</td>
<td>4.93</td>
<td>6.79</td>
<td>2.93</td>
</tr>
<tr>
<td>PROPERTIES &amp;</td>
<td>420</td>
<td>6.68</td>
<td>17.62</td>
<td>68.10</td>
<td>3.33</td>
<td>0.95</td>
<td>9.76</td>
<td>0.24</td>
</tr>
<tr>
<td>CHARACTERISTICS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHENOMENA</td>
<td>1441</td>
<td>22.92</td>
<td>24.43</td>
<td>63.43</td>
<td>4.44</td>
<td>3.05</td>
<td>4.51</td>
<td>0.14</td>
</tr>
<tr>
<td>MATERIALS</td>
<td>1186</td>
<td>18.87</td>
<td>54.13</td>
<td>37.69</td>
<td>5.14</td>
<td>1.69</td>
<td>1.26</td>
<td>0.08</td>
</tr>
<tr>
<td>INTELLECTUAL TOOLS</td>
<td>310</td>
<td>4.93</td>
<td>17.42</td>
<td>71.29</td>
<td>6.45</td>
<td>1.94</td>
<td>2.90</td>
<td>0</td>
</tr>
<tr>
<td>EQUIPMENT &amp; STRUCTURES</td>
<td>639</td>
<td>10.17</td>
<td>26.60</td>
<td>58.84</td>
<td>9.55</td>
<td>1.88</td>
<td>3.13</td>
<td>0</td>
</tr>
<tr>
<td>SCIENCES &amp; TECHNOLOGY</td>
<td>230</td>
<td>3.66</td>
<td>56.09</td>
<td>28.70</td>
<td>7.83</td>
<td>4.78</td>
<td>2.61</td>
<td>0</td>
</tr>
<tr>
<td>ORGANISMS</td>
<td>413</td>
<td>6.57</td>
<td>36.32</td>
<td>51.33</td>
<td>8.47</td>
<td>1.94</td>
<td>1.94</td>
<td>0</td>
</tr>
<tr>
<td>EARTH, SPACE &amp; TIME CONCEPTS</td>
<td>566</td>
<td>9.01</td>
<td>18.37</td>
<td>71.20</td>
<td>8.30</td>
<td>1.24</td>
<td>0.71</td>
<td>0.18</td>
</tr>
<tr>
<td>ORGANIZATIONAL, SOCIOECONOMIC, LEGAL, PERSONNEL ASPECTS</td>
<td>268</td>
<td>4.26</td>
<td>31.34</td>
<td>50.00</td>
<td>9.70</td>
<td>0.75</td>
<td>8.21</td>
<td>0</td>
</tr>
<tr>
<td>CONSTANTS</td>
<td>36</td>
<td>0.57</td>
<td>13.69</td>
<td>80.56</td>
<td>2.78</td>
<td>0</td>
<td>2.78</td>
<td>0</td>
</tr>
<tr>
<td>RESIDUAL</td>
<td>25</td>
<td>0.39</td>
<td>12.00</td>
<td>80.00</td>
<td>4.00</td>
<td>4.00</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CONCEPTS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Facets</td>
<td>6285</td>
<td>100.00</td>
<td>31.54</td>
<td>54.65</td>
<td>7.49</td>
<td>2.42</td>
<td>3.85</td>
<td>0.43</td>
</tr>
</tbody>
</table>
examples, ASFIS descriptors are listed first, followed by LC Subject Headings, and according to the typography employed in the two lists, ASFIS descriptors are all in upper case, while LC Subject Headings have only initial letters of each part in upper case.)

The similar matches fall into a number of categories. Type 1 is natural order (ASFIS) vs. inverted order (LCSH). Examples include: NATURAL GAS vs. Gas, Natural; UNDERWATER PHOTOGRAPHS vs. Photography, Submarine; and UNMANNED VEHICLES vs. Vehicles, Remotely piloted.

Type 2 is term (ASFIS) vs. term plus limiting words (LCSH): MUTATIONS vs. Mutations (Biology); SOCIAL BEHAVIOUR vs. Social behavior in animals; OUTFALLS vs. Ocean outfalls; and PREDATORS vs. Predatory animals.

Type 3 is a term (ASFIS) vs. a term plus a related term (LCSH): RIGGING vs. Masts and rigging; and CILIA vs. Cilia and ciliary motion.

Type 4 is initials (ASFIS) vs. words (LCSH): PCB vs. Polychlorinated biphenyls; and DNA vs. Deoxyribonucleic acid.

Type 5 is an adjective plus a term (ASFIS) vs. the term as a subdivision (LCSH): WATER FILTRATION vs. Water - Purification - Filtration; and AVIAN PHYSIOLOGY vs. Birds - Physiology.

Type 6 is a term vs. a term plus an explanatory adjective or noun. This difference goes both ways, thus: TANKER SHIPS vs. Tankers, but HYDROFOILS vs. Hydrofoil boats, and ICE BREAKERS vs. Ice-breaking vessels.

Type 7 simply involves different vocabulary: BRINES vs. Salt, and COMMUNICATION SATELLITES vs. Artificial satellites in telecommunication.

When comparing the “Operations” facet against LCSH, many terms register as ST, because LCSH tends to use instruments or agents rather than activities. Most of these LCSH equivalents have an exact match, however, in the “Equipment and Structures” or “Organisms” facets. Examples include FLOW MEASUREMENT and FLOWMETERS vs. only Flow-meters; and PHOTORECEPTION and PHOTORECEPTORS vs. only Photoreceptors.

The match category DT (different terms) applied to 2.4% of all ASFIS descriptors, with the highest number (49%) in the “Operations” facet, and the lowest (0%) in the “Constants” facet. These DT matches fall into at least four types. Type 1 involves technical terms vs. layman’s terms and goes both ways: HEAVY WATER vs. Deuterium oxide; IRONSTONE vs. Tiger’s eye; and EXPLORATORY BEHAVIOUR vs. Curiosity.

Type 2 is the discipline vs. the subject studied: CONCHOLOGY vs. Mollusks.
In type 3, different concepts are used, probably incorrectly, as equivalents: MICROHABITATS vs. Niche (Ecology), and AGROPISCICULTURE vs. Fish - Culture.

Type 4 can only be characterized as the use of different vocabulary for the same concept: OCEAN DUMPING vs. Waste disposal in the ocean; DISEASE RESISTANCE vs. Natural immunity; LEE WAVES vs. Mountain wave; and SINGLE POINT MOORINGS vs. Petroleum shipping terminals.

Of all ASFIS descriptors, 3.9% had a narrower term (NT) equivalent in LCSH, with the highest in “Properties and Characteristics” (9.8%) and the lowest in “Residual Concepts” (0%). Examples include: EXPLORATION vs. Underwater exploration; Outer space exploration; etc.; CIRCULATION vs. Lakes - Circulation; Pulmonary circulation; etc.; and CONVECTION vs. Convection (Meteorology); Convection (Astrophysics); etc. In the first two examples above, although the term appears in the ASFIS Thesaurus, the scope note recommends using a more specific term, which would lead to a “+” match in this study.

The final match category BT (broader term equivalents in LCSH) accounted for only 0.43% of all ASFIS descriptors, with 2.9% in “Operations” and none or nearly none in most other facets. Examples include ARTISANAL FISHING vs. Fishing; and RIVER VALLEYS vs. Valleys.

DISCUSSION

The overall match rate of ASFIS descriptors with LCSH is very low. Several other comparisons of subject-specific thesauri and encyclopedias with LCSH have showed a similar poor match rate, with LCSH in all cases “...found to be deficient when measured against other bibliographic standards in the field.” (4)

This study compares subject headings and descriptors. As noted by Cochrane (5), the fine distinctions escape most people, but certain basic differences are easily recognized. Subject headings are at least 100 years old (6), while descriptors date back to about 1959 (7). The sources for subject headings and descriptors differ. LC Subject Headings are based on literary warrant, that is, terms are selected from what appears in the literature. “Systems based on literary warrant grow by accretion for the most part. As time passes, logic and consistency suffer...” (8). Despite the principle of literary warrant, many terms with obvious literary warrant are lacking in LCSH, or are only added very late (4). The other approach is followed by designers of thesauri. The ASFIS terminology is based on related authority lists, other thesauri, and glossaries (1). In contrast to LCSH, which reacts to literature, the ASFIS Thesaurus provides a vocabulary to be used and influences the terminology used in literature.

Boll (6) discussed a number of differences between subject headings and descriptors and noted that many of the traditional differences are breaking down. Differences
relevant to this study include the following: 1. Subject headings are applied by librarians, whereas descriptors are applied by non-librarians to materials purchased and used by librarians, e.g. the ASFA database. 2. One subject heading summarizing the entire content of the work is to be assigned, whereas enough descriptors are assigned to cover all important aspects of a work. 3. A corollary of this is that minor topics do not receive subject headings, whereas primary and minor topics receive descriptors. 4. Subject headings are often inverted, reflecting "...a reluctance to disperse related entries" (3), whereas multiword descriptors are used in natural word order. 5. Subject headings are pre-coordinated, whereas descriptors are post-coordinated. 6. LC Subject Headings have subdivisions, one form of pre-coordination, whereas descriptors are not subdivided. This explains most of the cases where the LCSH equivalent to an ASFIS descriptor was a narrower term. In most of these cases, as noted above, the ASFIS Thesaurus scope note recommends using a narrower term, which is also listed. 7. Subject headings are applied on the principle of economy of input and redundancy of search, whereas descriptors emphasize redundancy of input and economy of search. These various differences have been a source of complaints, mostly against LCSH. O'Neill and Aluri (9) note that the "summative approach of subject analysis ... severely limits users' ability to ... retrieve records by subject." Many have recommended that LCSH become more like a thesaurus (5).

The differences are breaking down, however (6). Many descriptors have been pre-coordinated to solve the problems of uncontrolled natural language (7), thus making them more like subject headings. In general, the trend is for LC Subject Headings to be applied more in the manner of descriptors. One analysis showed that the average number of subject headings applied to all LC records is 1.4, but for science it is 1.87 and for technology 1.6, or viewed as percentages, 33% of science/technology records have 2 subject headings and 17.8% of science and 15% of technology records have 3 subject headings (9). This is still far from the many descriptors applied to each work, but departs from the recommended single subject heading per record.

The advent of online searching has affected the way in which subject headings can be utilized. Bibliographic utilities such as the Research Libraries Information Network (RLIN) which have a capability for Boolean subject searching, combined with truncation, allow post-coordination of subject headings and uncoupling of subdivisions, thus allowing much more precise searching. There are apparently no figures on this, but it would seem that catalogers using facilities such as RLIN may be assigning more subject headings to items they catalog, that is, there is more redundancy of input, leading to economy of searching and allowing more post-coordination. Of the differences listed above, the inversion of LCSH does not seem to be changing, although this has often been recommended. Online searching makes inversion unnecessary, since descriptors or subject headings can be gathered together through post-coordination. Thus, many of the weaknesses of LCSH may be overcome by online searching.
There remains the basic problem of choice of terminology. The ASFIS Thesaurus fits aquatic science much more precisely than does LCSH. The same is undoubtedly true of thesauri in other fields (4). The ASFIS Thesaurus tends to influence and standardize terminology used in literature. LCSH reacts, although often slowly, to what appears in the literature. It seems that LCSH could be greatly improved, and more up to date, if the compilers of LCSH relied on professionally accepted thesauri, such as the ASFIS Thesaurus, as the source of terminology for a particular field, rather than attempting to deduce terminology from the literature.

REFERENCES


