Available online at www.sciencedirect.com
SCIENCE DIRECT®
Information and Software Technology 47 (2005) 245–255
INFORMATION AND SOFTWARE TECHNOLOGY
www.elsevier.com/locate/infsof

A unified classification system for research in the computing disciplines

Iris Vessey[a,*], V. Ramesh[a], Robert L. Glass[b]

[a]Department of Accounting and Information Systems, Kelley School of Business, Bloomington, IN 47405, USA
[b]Computing Trends, Bloomington, IN, USA

Received 31 December 2003; revised 22 July 2004; accepted 23 August 2004
Available online 2 October 2004

Abstract

The field of computing is made up of several disciplines of which Computer Science, Software Engineering, and Information Systems are arguably three of the primary ones. Despite the fact that each discipline has a specific focus, there is also considerable overlap. Knowledge sharing, however, is becoming increasingly difficult as the body of knowledge in each discipline increases and specialization results. For effective knowledge sharing, it is therefore important to have a unified classification system by means of which the bodies of knowledge that constitute the field may be compared and contrasted. This paper presents a multi-faceted system based on five research-focused characteristics: topic, approach, method, unit of analysis, and reference discipline. The classification system was designed based on the requirements for effective classification systems, and was then used to investigate these five characteristics of research in the computing field.

Keywords: Classification system; Computing; Research method; Research approach; Reference discipline; Unit of analysis

1. Introduction

Computer Science (CS), Software Engineering (SE), and Information Systems (IS) share a number of areas of interest. CS is the scientific discipline concerned with the fundamentals of computing. SE applies those fundamentals to the development of software systems. IS fulfills the computing, and especially information, needs of the business community. Thus these disciplines have certain elements in common—computing concepts, systems development, and information technology—but they also have clearly distinguishable goals.

Because of the interrelated nature of the three disciplines, it is important that they share knowledge. The need for such sharing is apparent in all colleges and universities offering computing subject matter, but is especially apparent in those schools with integrated schools of computing. Further, knowledge sharing is becoming increasingly difficult as the body of knowledge in each discipline increases and specialization results.

To facilitate knowledge sharing, disciplines typically develop classification systems, which then provide a common terminology for communication. Frequently, the classification system, which most often is based on topic, is adopted by a society or a journal or journals and the entries in the system are used as index terms and/or keywords to characterize material published in those journals, as well as in books and other types of literature. For example, the ACM has defined a set of index terms to be used by its journal, Computing Reviews, and other ACM journals require authors to use the same terms to characterize submitted articles. To support research endeavors, the index terms or keywords can then be used to identify articles using computerized information retrieval systems, for example.

Because knowledge in the three computing-related disciplines may be found in any of the contributing disciplines, it is important that the field (of computing) as a whole have a classification system that enables the identification of relevant research across the disciplines. The thesis of this research is that existing classification systems do not fulfill these needs. The objective of this research is therefore to develop and present a unified classification system for characterizing computing research.

* Corresponding author. Tel.: +1 812 855 3485; fax: +1 812 856 4985.
E-mail address: ivessey@indiana.edu (I. Vessey).

0950-5849/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.infsof.2004.08.006

The paper proceeds as follows. Section 2 examines the predominant classification systems in each of the disciplines with the objective of assessing their effectiveness in providing a unified view of the field. This analysis is followed by a discussion of the characteristics of effective classification systems. The paper then presents a unified classification system designed to characterize research in all three computing disciplines and to meet the requirements for effective classification systems. The system is comprehensive in that it addresses all of the characteristics currently identified as being important in distinguishing research of this nature. We then describe how we tested the classification system by classifying approximately 1500 research articles using the system. The paper concludes with a discussion of the implications of using such a classification system to classify research in the computing disciplines.

2. Existing computing classification systems

Here we examine the predominant classification systems currently used in CS, SE, and IS to determine their effectiveness in providing a unified view of the field. By predominant, we mean the classifications that are intended to be representative of the discipline rather than focusing on a more limited sub-set of relevant research. Note, also, that the classification systems selected for examination within each discipline are those most commonly used to specify keywords in research articles.

2.1. Computer Science

Computer Science does not have a widely-accepted classification scheme. However, CS research published in ACM journals and magazines uses the categories defined in the ACM Computing Classification System to identify index terms [1]. Fig. 1a presents the top two levels of this classification system, which has a total of four levels. There are 11 top-level categories and 81 second-level categories with total entries approaching 1400.

This classification system focuses on topics within the computing disciplines. Note that both IS and SE appear in the classification, IS at the highest level (H), indicating that it is a subset of CS, and SE at the second highest level (D.2), indicating that it is a sub-set of Software.

Perusal of the classification system reveals that the categories may not support unambiguous classification. For example, K.6 Management of Computing and Information Systems under K. Computer Milieux refers to both Software Engineering (entry D.2 cross-referenced to category K.6.2) and Information Systems (by name). These observations indicate that the CS classification system does recognize that SE and IS contribute to the 'CS' body of knowledge. What is equally clear, however, is that it does not represent them effectively. Further, many interests of SE, such as Computer-Communication Networks (C.2), Performance of Systems (C.4), Computer System Implementation (C.5), Programming Techniques (D.1), and Programming Languages (D.3), do not appear under the topic of Software Engineering. Similarly, IS is represented not only within in its own topic, but also under J.1 Administrative Data Processing and E.5 Files. Further, E.4 Coding and Information Theory is cross-referenced to H.1 Models and Principles (of Information Systems). In addition, many areas of interest to IS researchers such as, adoption and diffusion of technology, organizational learning and strategy, etc. do not appear in the classification system.

These observations suggest that using the ACM Computing Classification System does not lead to the unambiguous designation of categories and, further, that it does not allow proper characterization of the computing disciplines other than computer science.

2.2. Software engineering

Many of the leading SE publications record keywords or index terms with each research paper However, they provide little or no direction about how to derive them. IEEE Transactions on Software Engineering, for example, states [19]: 'Index terms should be relatively independent, and as a group should optimally characterize the paper.' The 'Guide for Authors' for Information and Software Technology [20] states that the authors should provide 'one to five keywords,' goes on to say that 'keywords are essential,' and provides consistency guidelines for their choice. Finally, the Journal of Systems and Software does not request keywords or index terms in its Information for Authors; however, it does require the author(s) to provide keywords at the time of proof reading the about-to-be-published paper.

Whether software engineers develop their keywords based on the ACM Computing Classification System, papers in other SE journals, or their own judgment is not known. What is clear, however, is that SE does not have a classification system to aid in organizing and therefore communicating its body of knowledge.

2.3. Information systems

The IS discipline, per se, has no widely-accepted classification(s) to reflect the characteristics of interest. However, in 1988, one of the top journals in the discipline, Management Information Systems Quarterly (MISQ), published an academic paper on a 'keyword classification system' for characterizing IS research, followed by an update in 1993 [6,7]. Following the publication of the first of these articles, MISQ has until recently, required that articles submitted to the journal include keywords based on this classification system. Its use has not, however, spread to other IS journals.

The top two levels of the classification system, which has a total of four levels, are presented in Fig. 1b. There are nine