Interdisciplinarity and the Classification of Scholarly Documents by Phenomena, Theories, and Methods

Rick Szostak

University of Alberta, Department of Economics, Edmonton, Canada. rick.szostak@ualberta.ca
Visiting Fellow, European University Institute, Department of History and Civilization, Firenze, Italy.

Abstract
The paper argues that information science can best serve the needs of interdisciplinary scholarship (which is of increasing importance) by developing universal classifications of the phenomena studied by scholars and the theories and methods applied by scholars. Present systems of document classification are grounded in disciplinary terminology and thus serve interdisciplinary scholarship poorly. The second part of the paper outlines the importance of the recommended type of system of classification, the limitations of present systems, and the effects of these limitations on interdisciplinary scholarship. The third part argues that such a system of classification is feasible, and that it is best developed through a combination of induction and deduction.

Keywords: Classification, Interdisciplinarity, Methods, Phenomena, Theories.

Resumen
La comunicación argumenta que la Ciencia de la Información puede servir mejor a las necesidades de los estudios interdisciplinares (cuya importancia es creciente) desarrollando clasificaciones universales de los fenómenos, de las teorías y de los métodos que utilizan los investigadores. Los actuales sistemas de clasificación de documentos se fundamentan en terminología disciplinar, de ahí que resulten limitados para acometer estudios interdisciplinares. La segunda parte de la comunicación señala la importancia del tipo de sistema de clasificación recomendado, las limitaciones de los sistemas actualmente en vigor y los efectos de estas limitaciones en los estudios interdisciplinares. La tercera parte manifiesta que el sistema de clasificación propuesto es viable, y que para su adecuado desarrollo conviene utilizar una combinación de métodos inductivos y deductivos.

Palabras clave: Clasificaciones, Fenómenos, Interdisciplinariedad, Métodos, Teorías.

1 The Big Picture

The main hypothesis of this paper is that: *The feasible strategy for information science that would most benefit interdisciplinary scholarship is the classification of scholarly documents in terms of universal classifications of phenomena, theories, and methods*. This hypothesis follows in turn from the motivation for this conference: that interdisciplinary scholarship is of increasing importance in the academy, but could be better served by information science. As Palmer (2001, xii) has noted, if it is useful to enhance the productivity of scholarship, then it is valuable to enhance access to scholarly information, for this lies at the heart of the scholarly project.

This hypothesis suggests that such a classification is both desirable and feasible. Both points need to be firmly established before information scientists are likely to engage in the development of such a classification. However, once these points are established, the rising importance of interdisciplinarity, in concert with the digitization of document classification, combines to generate a historical moment in which the development of a radically new system of document classification may be pursued.

This paper strives to combine in one place the key arguments regarding both the desirability and feasibility of the recommended system of classification. In both instances it seeks to address potential or actual objections.

2 The Desirability of the Classification

To argue that a new system of classification would better serve interdisciplinary scholarship must involve both establishing the advantages of the new system and the limitations of the old. It must first be established then that searching in terms of phenomena, theories, and methods is of critical importance to interdisciplinarians. It must then be shown that existing systems of document classification serve these needs poorly (theories and methods are first addressed, and then phenomena). Finally, the effects of these limitations on interdisciplinary scholarship are addressed.

2.1 The Importance of Phenomena, Theories, and Methods

Weinberg (1988) famously noted that researchers in general are poorly served by classifications of documents solely in terms of what these are 'about': novices search for books 'about' a particular topic, but scholars seek works that express certain 'ideas.' They seek works that apply particular scholarly perspectives (Weinberg stressed theories but addressed methods) to particular subjects. Palmer (1996) also urges the classification of documents in terms of theory and method applied.

Scholars of interdisciplinarity have independently reached the same conclusion as Weinberg regarding the particular importance of classification by theory, method, and phenomena. Both Klein (1990) and Salter and Hearn (1996) identify these as key identifying characteristics of disciplines, and thus the key elements that interdisciplinary scholarship must integrate across. To be sure, these authors speak of two other disciplinary characteristics: disciplinary perspective or worldview, and disciplinary rules. The former involves the general approach (including ideological, epistemological, and ethical elements) of members of particular

disciplines, and the latter involves the standards of peer evaluation that characterize a discipline. Notably, in practice, both of these primarily serve the interdisciplinarian in evaluating how the insights produced by a disciplines theories and methods may have been shaped by its worldview and incentive structure (see Repko 2006). For present purposes, it is most useful to note that these characteristics operate at the level of disciplines rather than individual works, and are thus best mastered by reading general works about disciplines or interdisciplinarity.

Szostak (2003b) summarizes the analyses of Klein and Salter and Hearn, and develops the following (partial) definition of interdisciplinarity: 'Openness to the application of all theories and all methods to any set of phenomena.' In other words, interdisciplinary practice generally involves the integration of theories and/or methods from multiple disciplines, and/or the study of relationships between phenomena usually studied by different disciplines.

Special attention should be paid here to scholarly concepts. Hjørland and Nissen Pedersen (2005, 586) follow Rey (1998, 505) in arguing that '… there is considerable disagreement about what exactly a concept is' but that concepts '… seem essential to categorizing the world.' Hjørland and Nissen Pedersen thus advocate an entirely inductive approach to classification whereby concepts are identified in use. At least with respect to scientific documents, though, it would seem that concepts should refer to specific phenomena or sets of phenomena, or theories or methods or components thereof (Wallace and Wolf, 2006, 4-5; Repko 2006; Szostak 2007), since these are the key elements of scholarly discourse. It may well be that concepts that cannot be defined in these ways usually do not serve to advance scholarly understanding. For present purposes, though, the key point is that the slippery concept of scholarly concepts can most often be handled by careful classification of phenomena, theories and methods.

2.2 The current approach to theory and method

As noted above, present classification systems rarely classify documents with respect to the theories and methods employed. Since they focus on what works are about, only works about theories and methods are classified in terms of the theory or method in question. Notably, classifications thus treat theories or methods as if they were phenomena.

Even the classification of works about theories can be problematic. Hjørland and Nissen Pedersen (2005) report on three quite distinct uses of the phrase 'activity theory': different disciplines intend quite different causal processes by the same nomenclature. Even more troubling are cases where quite similar theories or techniques go by quite different names in different disciplines. In such cases, researchers cannot readily identify all relevant works about a particular theory or technique.

Weinberg noted that the limitations of classifications forced scholars to rely on time-consuming and imperfect search strategies such as asking colleagues or following citation trails (the strategy recommended by Bates, 1996, and Palmer, 1996). Searches that crossed disciplinary boundaries were particularly likely to miss important works.

2.3 The current approach to phenomena

Present classification systems are primarily discipline-based, and thus the same phenomenon is classified multiple times and often treated differently in different disciplines. As Bulick described as early as 1982, this approach has caused great confusion as disciplinary

boundaries have shifted and interdisciplinary fields have emerged. Three broad types of problem emerge: phenomena that are studied by more than one discipline receive multiple classifications; subjects that are inherently interdisciplinary have no obvious place; and subjects that combine existing subjects have no obvious place. These problems violate one of the fundamentals of any system of classification: that any document should have one and only one place in the system. Yet for interdisciplinary researchers these problems alone comprise an inconvenience rather than a barrier to research. They force researchers to look in multiple places for related works (and thus make 'browsing' the shelves a more challenging enterprise, and increase the likelihood that relevant works will be missed). But as long as the researcher knows what subject to search under they will eventually find the place or places where such works are located.

Far more troubling is the fact that quite different terminology is used in different disciplines. As with theories and methods (above), the researcher will miss relevant works if they do not know what terms to search for. They may fall back on general works about disciplines, but this is a time-consuming strategy for identifying terminology. Moreover it presumes that they know at the outset which disciplines to investigate. Yet one of the challenges of interdisciplinary research is to identify relevant disciplines (Repko, 2006). And the most useful information is often the most surprising, and this will usually be information the researcher would not have searched for (Palmer, 2001).

As noted above, interdisciplinary research often examines links between phenomena investigated by different disciplines. The interdisciplinary researcher must first identify the set of relevant phenomena, and this task will be particularly difficult in the absence of a universal classification of phenomena.

While the academy benefits from the complementary efforts of specialized and interdisciplinary researchers, even specialized researchers can benefit from familiarity with related work in other disciplines: this will not only suggest new avenues of research but remind them of the biases that may affect their disciplinary approach (Szostak 2004). Such knowledge would be much more likely if works on the same topic from different disciplines were classified and shelved together.

2.4 Some Effects of the Present Situation

We have already had cause in the preceding sections to note some of the malign effects on interdisciplinary research of the limitations of present systems of document classification. Put simply, interdisciplinary researchers are not facilitated as they could be in their search for related material in different disciplines, whether the relatedness stems from similarities in phenomena studied and/or method or theory applied. This section simply lists some questions that should occur to the interdisciplinary (and often the disciplinary) researcher, but are unnecessarily difficult to answer within the present state of affairs:

- What theories and methods have been applied to the study of a particular set of phenomena in the past?
- To what set of phenomena has a particular theory or method been applied? (Note that one of the key tasks of scholarship is identifying the range of applicability of theories.)
- What problems have been encountered in these endeavors? (This question cannot be entertained until the more basic questions are answered.)

3   The Feasibility of the Classification

As noted at the outset, it is necessary to establish both desirability and feasibility. A desirable change that is thought to be impracticable will not be pursued. The fact that documents have not been classified in the past in terms of theory and method applied, or in terms of a universal classification of phenomena studied, suggests that these enterprises might be difficult. If classifications in terms of theory and method are so desirable, why otherwise have they not been attempted? Weinberg (1988) herself, despite her concerns, made no recommendations for change, feeling that efforts to classify in terms of theory and method applied would be too complicated. Gnoli, Bosch, and Mazzocchi elsewhere in this volume discuss the efforts by members of the Classification Research Group (especially Douglas Foskett and Barbara Kyle) to develop universal classifications of phenomena; why if feasible were these efforts not sustained by other researchers?

One part of the answer to these questions is simply that the major systems of cataloguing documents were each developed during the age of card catalogues. It made sense in that environment to focus on classification in terms of subject matter. As the movement to classify a wider range of meta-data attests, the digitization of library classifications encourages the classification of documents along a wider range of dimensions.

A second part of the answer reflects the need for interdisciplinarity in document classification itself. Both classificationist and classifier need little knowledge of the content of the documents they address if they will classify these only in terms of how the documents fit into a narrow disciplinary conception of subject matter. If instead these will be classified in terms of a universal classification of phenomena, and in terms of theory and method applied, much more careful reading of texts – or collaboration between information scientists and other scholars – will be called for. The inductive approach advocated by Hjørland and Nissen Pedersen (2005) and others represents one promising strategy for developing better classifications based on more careful analysis of texts. But as those authors freely admit a universal classification is unattainable with their approach. A more deductive approach (which can complement the inductive approach) is necessary (Szostak 2007). Especially with respect to the classification of theories and methods this is likely to require collaboration between scholars of scientific practice and information scientists.

3.1 Classifying with respect to method applied

This may well be the easiest element of the proposed system to achieve. Broadly speaking, there are only a dozen scholarly methods: classification itself, experiments, interviews, surveys, observation, statistical analysis, mathematical modeling, textual analysis, mapmaking, hermeneutics/semiotics (the study of symbols), the study of physical traces (as in archaeology), and experience/intuition (see Szostak 2004, 101-2). These can be disaggregated into a manageable set of more specific tools and techniques. These are often (though not always) referred to by similar terminology in different disciplines: the terminology of econometrics infuses statistical analysis in other disciplines.

It should be stressed that the classification of a document in terms of method applied will be independent of its classification in terms of phenomena studied (which will determine its shelf placement) and theory applied. The notation for these three distinct elements can be part of analytico-synthetic and expressive class marks. Researchers using electronic catalogues will be able then to search by combinations of theory, method, and phenomena.

3.2 Classifying with respect to theory applied

Theories, alternatively, present great challenges for classification. One problem is that there are a potentially infinite number of theories, for new theories are invented every day in the scholarly enterprise. On its own this is a challenge often faced in classification, and one that can be dealt with by a hospitable (expandable) classification with room for new classes.

More importantly, and as noted above, there is considerable terminological confusion in the area of theory: the same title encompasses quite different types of theory, while quite similar theories go by different names. Even advocates of a particular theory quarrel about what the essence of their theory is (see, for example, Turner, 2000). A classification grounded only in the names used by authors to describe their theories would thus generate the same sorts of confusion that classification of subject matter in terms of disciplinary categories does. Nor is the only problem one of cross-disciplinary communication, for theories evolve through time (much more than do definitions of phenomena, though these if not carefully defined may evolve too), and thus an author using a theory today may have quite different ideas from an author using the same theory a decade before. While a classification in terms of the theoretical terms used by authors would be an important improvement over no classification by theory whatsoever, it would be inappropriate in important respects.

The terminological problem might be approached inductively. Recall, though, that inductive methods are generally applied only within particular scholarly communities. Happily, a deductive approach is available. A variety of theory types can be identified along several key dimensions: who is the active agent(s) in the theory, what do they do, why do they do it, what sort of process through time is envisaged, and how generalizable is the theory? (see Szostak 2003a, 2004 for details). [Note that this set of dimensions, though logically derived, is expandable should other useful dimensions be identified.] Individual theories can be placed within the typology, and thus works can be classified simultaneously by theory and by theory type. Researchers can search by theory name or by a particular type of theory.

This approach solves immediately the problem of different theories operating under the same name. The catalogue will inform the researcher searching by theory name that the particular name is associated with quite different theories. And the researcher can quickly investigate other theories similar to the researcher's own theory in certain respects: same type of agency or action or dynamic process.

As noted above, theories evolve through time. And theorists disagree about what a particular theory entails. As a result many theories sprawl across multiple theory types. While this fact complicates the life of the classifier – they must look beyond theory name – it has enormous advantages for the researcher who may wish to only read 'Marxian theory of a particular type.'

Several grand theories were placed within the typology of theory types in Szostak (2003d) and (2004). The same strategy should prove even easier for narrower theories.

3.3 Classifying with respect to phenomena

The (until recently) independent efforts of Szostak – drawing on the study of science literature – and Gnoli and colleagues (see his co-authored paper in this volume, Gnoli and Poli 2004, and especially the website at www.iskoi.org/ilc/how.htm) – drawing on the

information science literature – have produced similar and entirely complementary approaches to the development of a universal classification of phenomena. While these schemes are in their early stages, the broad outlines are clear, and efforts to classify some literatures have been successfully undertaken (see especially the ISKO Italy website noted just above). Though not itself an effort at document classification, Szostak (2003c) established that *the arguments of* hundreds of works from across the human sciences could be classified in terms of a simple but universal classification of phenomena. While much more remains to be done, enough has been accomplished to establish the feasibility of the endeavor.

As with the earlier works of the CRG, these classifications are hierarchical. A manageable set of main classes of phenomena is identified, and these are unpacked or disaggregated into multiple levels of constituent phenomena. (They also notably each utilize a freely faceted approach).

Importantly, these phenomena can usually be precisely defined, either in terms of their internal essence or their function. Gnoli and Poli (2004) argue that functional definitions only make sense for phenomena of a level of complexity equal to or beyond biological organisms; for atomic particles and chemical elements definition in terms of essence are superior. It can thus be expected that a universal classification of phenomena will not be invalidated by time: the phenomena included will retain their meaning.

On the frontiers of some fields of scientific research – notably the physics of sub-atomic particles – phenomena are defined provisionally with respect to certain theories. These phenomena are observed to change definitionally as the theory is clarified. In some cases, it may be best to treat such phenomena as aspects of the theories that suggest their existence until scholarly consensus on theory is achieved. In general, though, the phenomena studied by scholars can and should be defined without any reference to theories.

Most scholarly research studies not just one phenomenon but the influence that one or more phenomena exert on others. A scholar interested in how, say, economic productivity is influenced by the level of trust in a society will wish to consult works investigating that 'causal link' but not every work on trust or economic productivity. This outcome can be achieved by using linked notation between phenomena to indicate the main causal link(s) pursued in a work. Existing systems of classification generally classify works in terms of only one subject. This focus on causal links would further enhance the utility of document classifications to researchers. Note in this regard that discipline-based classifications would face difficulty in capturing interdisciplinary causal links. The different terminology used for each phenomenon involved would multiply the number of ways in which the same link might be classified.

3.4 Induction and deduction

The various classifications in Szostak (2004), as well as the classification of Gnoli and colleagues, were developed through a mix of induction and deduction. Surveys of the literature suggested key classes. Logical classifications were developed and then adjusted as applied further to the scholarly literature. The informal inductive procedure employed could usefully be complemented by the more formalized inductive techniques advocated by authors such as Hjørland and Nissen Pedersen (2005).

Hjørland and Nissen Pedersen (2005) argue that their inductive approach is only suited to the classification of the literatures of individual scholarly communities. This paper suggests that this need not be the case. To be sure, a greater variation in meanings will be found as one searches a wider sample of literature. Potentially at least, large enough differences in meaning could result in different classifications. Notably, when López-Huertas investigates the interdisciplinary field of gender studies inductively, she finds that roughly half of the concepts employed are borrowed from other disciplines 'with the same form and apparently the same sense as they have in their original realm' (2006, p. 333). The 32% of terms developed independently within gender studies, and especially the remainder that are developed through interaction with other disciplines, may also share similar meanings in different realms.

Hjørland and Nissen Pedersen (2005) are guided, as is much recent research in information science, by a laudable desire to reflect the differences in perspective that characterize modern scholarship. Yet their strategy serves to create barriers between communities with different perspectives. The scholarly project can only benefit from different worldviews if there is communication between communities. The approach recommended in this paper is also grounded in the recognition of different scholarly perspectives, but allows these to be more readily juxtaposed and thus integrated.

It has been suggested (but not shown) that the most ambiguous scholarly concepts are those which do not represent some element of phenomena, theory, or method. By focusing our efforts on phenomena, theory, and method, we identify the core elements of a work, and simultaneously those elements where language ambiguity is the least. And by wedding inductive to deductive analysis we are able to further clarify the precise meanings of individual phenomena, theories, and methods.

It should also be stressed that most works within a discipline can readily be classified in terms of phenomena, theory, and method. For example, a book on the sociology of culture could be classified in terms of the particular cultural elements studied and the theories and methods applied. Only general treatments of sociology as an enterprise would need to be classified as sociology (a phenomenon in the class of types of scholarship/science).

4   Concluding remarks

This paper is in some sense a manifesto for a radically new approach to document classification. It is both highly desirable and feasible to classify scholarly documents in terms of a universal classification of phenomena, theories and theory types, and methods. If information scientists develop the sort of classifications suggested above, they will greatly facilitate interdisciplinary scholarship. Many scholars at present understand the value of interdisciplinary scholarship but hesitate to engage in this because of its challenges. Given that specialized and interdisciplinary scholarship are mutually supportive, greater efforts toward interdisciplinarity will markedly enhance the quality and productivity of the scholarly enterprise as a whole. In other words, information scientists can at this historical moment have a huge and beneficial impact on the future course of scholarship by developing classifications that facilitate interdisciplinary analysis.

References

BATES, M. J. Learning about the information seeking of interdisciplinary scholars and students. *Library Trends,* 1996, vol. 45, n.2, p. 155-64.

BULICK, S. *Structure and subject interaction.* New York: Marcel Dekker, 1982.

GNOLI, C.; POLI, R. Levels of reality and levels of representation. *Knowledge Organization*, 2004, vol. 31, n.3, p. 151-60.

HJØRLAND, B.; NISSEN PEDERSEN, K. N. A substantive theory of classification for information retrieval. *Journal of Documentation,* 2005, vol. 61, n. 5, p. 582-95.

ISKO ITALY 2004. *Integrative level classification: research project* [electronic resource]. <www.iskoi.org/ilc/>. [Consulted: 15 nov. 2006]

KLEIN, J. T. *Interdisciplinarity.* Detroit: Wayne State University Press, 1990.

LÓPEZ-HUERTAS, M. Thematic map of interdisc-iplinary domains based on their terminological representation: gender studies. In: International ISKO Conference (9th: Vienna: 2006). *Knowledge organization for a global learning society*: *proceedings.* Würzburg: Ergon, 2006, p. 331-8.

PALMER, C. L. Information work at the boundaries of science: linking library services to research practices. *Library Trends*, 1996, vol. 45, n. 2, p. 165-191.

PALMER, C. L. *Work at the boundaries of science*: *information and the interdisciplinary research process*. Dordrecht: Kluwer, 2001.

REPKO, A. *Interdisciplinary practice: a student guide to research and writing.* Manuscript under review, 2006.

REY, G. Concepts. In: CRAIG, E. (ed.). *Routledge Encyclopedia of Philosophy,* 1998, vol. 2, p. 505-17.

SALTER, L.; HEARN, A. *Outside the lines*: *issues in interdisciplinary research.* Montreal: McGill-Queen's University Press, 1996.

SZOSTAK, R. Classifying scholarly theories and methods. *Knowledge Organization,* 2003a, vol. 30, n.1, p. 20-35.

SZOSTAK, R. 'Comprehensive' curricular reform: providing students with an overview of the scholarly enterprise. *Journal of General Education,* 2003b, vol. 52, n. 1, p. 27-49.

SZOSTAK, R. *A schema for unifying human science*: *interdisciplinary perspectives on culture*. Selinsgrove PA: Susquehanna University Press, 2003c.

SZOSTAK, R. *Classifying science*: *phenomena, data, theory, method, practice*. Dordrecht: Springer, 2004.

SZOSTAK, R. Classification, interdisciplinarity, and the study of science, forthcoming. *Journal of Documentation,* 2007.

TURNER, B. S. (ed.). *The Blackwell Companion to social theory.* Oxford: Blackwell, 2000.

WALLACE, R. A.; WOLF, A. *Contemporary sociological theory: expanding the classical tradition,* 2006.

WEINBERG, B. H. Why indexing fails the researcher. *The Indexer,* 1988, vol. 16, n.1, p. 3-6.