# What kind of science should information science be?[1]

Belver C. Griffith

I propose today to argue that information science is a 'soft' science, that it should emphasize certain broad orientations in content and methodology, and that it faces certain perils and has certain opportunities in the near future in dealing with both its intellectual (content) problem and its institutional (resources) problem. I will first take a few moments to discuss my approach to the general problem of considering what kind of science information science 'should be'.

To state what a science 'should be' requires a critical, perhaps overwhelmingly difficult prior step, namely the creation of a system, or even some fragmentary ideas, that can distinguish among and characterize sciences. One cannot say what subject x 'should be' until one has the language resources for saying what subject x is. Unfortunately, the intellectual recourses available for discussing science are primitive - but not negligible. Among such resources are Merton's work on scientific norms; Price's continua of hardness and softness in science and his attempts to distinguish among sciences, technology and nonscience; Ben-David's ideas on institutionalization and internal direction of disciplines; Ziman's concept of scientific consensus and Ravetz's concept of scientific craft.[2] A broad approach, incorporating such conceptual material is emerging in the history of science. For example, scholars are beginning to generate quite a detailed picture of the intellectual context within which the great work of quantum mechanics emerged.[3] However, theirs is an art of retrospection - whereas mine must be one of prediction and prescription.

---

[1] This address is dedicated to all students who make professors work hard and especially to Sandra L. Dey and A. Ben Wagner.
To be published in the Journal of the American Society for Information Science.

This Spring I had the good fortune to conduct - with this address in the back of my mind - a seminar on information and knowledge. Throughout that seminar, I struggled with trying to discover or invent a general model for information science, and thereby, a general scheme for this paper. However, I felt like a swimmer, swept downstream by the current, trying to map a river. Accordingly, the address will deal with specific issues, and in particular, with try to identify appropriate matches, or their absence, among intellectual goals, people and content, and I believe that the philosophers and social scientists of science have given us some minimal tools to look at the inter-relationships of intellectual achievements, people and subject matter. To return to the image of a swimmer - I hope a few of you see some of the same rocks and embankments that I see.

In thinking on the broad issue, I have had to conclude that information science must be regarded as a 'soft' as opposed to 'hard' science. There is no particular clear definition of 'soft' and 'hard', but most people feel comfortable with contrasting the social 'soft' science with physical 'hard' science.[4] Within disciplines there are degrees of softness: experimental psychology is hard for psychology; molecular genetics is hard for biology.

Softness is attributed to fields that permit multiple approaches and mutiple answers. The subject matter in these fields is elusive and complex, and differences in points of view are necessary and expected. It is difficult to assess work in a soft field, no way in which to evaluate a discovery, so there are no outbursts of activity, as would follow a major finding in physics.

On first consideration, several features of information science that might suggest a degree of 'hardness'. They include:

- dependence on hard technology

This seems to me a relationship that does affect the internal structure of information science. At the principal interface, computer science, the principal commodity exchanged is a custom-made technology.

- training and background of 'pioneers'
- use of analytical tools, quantification

The first of these is superficial, a historical accident like the fact that Gestalt psychology was founded by three persons holding doctorates in physics. The second reflects the field's inability to avoid large numbers, a point to which I shall return.

- the 'cleaning' up of a point in theory or data

The most telling of these arguments is the fact that occasionally one can, regarding a particular issue, do a clearly better job than the next fellow, that is, collect better data, clarify a relationship, etc. However, the issue is almost always, to my view, merely a minor aspect of several major and different approaches to the main problem. Information science is 'soft' because people with all their glorious inflexibility and petty fickleness, are intrinsic to the systems we study whether as users, producers, operators, decision makers or as the ultimate, unanalyzable language processing device. Immediately, this introduces the possibility of a variety of approaches; theories are not clearly right or wrong; results are rarely the final answer. Both research problems and concepts linger on and on, as the literature of sociology will make clear.

Living with a 'soft' science is a difficult art at best and may be particularly difficult at the present moment for information science with its present personnel. In the United Kingdom and United States, the beginnings of information science were marked in the 1950's by the migration of a number of natural scientists, behavioural scientists, and applied mathematicians into problem areas dealing with information. The vast majority were, however, natural scientists, and despite the great contributions of these pioneers, there

have been major mismatches of approaches and subject matter.

Saracevic's excellent review of research on document relevance shows the way in which an apparently simple problem gradually is recognized to be a very complex one.[5] Why shouldn't it be a simple matter to partition a collection into relevant and irrelevant documents and determine how many of each are retrieved? Because the concept of 'relevance' involves one with all the complexities of human judgement and behaviour. Rather than furnishing simple answers, as apparently anticipated, the research ultimately showed that the problem was more like the experimental analysis of aesthetics. More recently, the science policy area has stumbled upon an issue held in common with us, namely, assessing the quality of documents and authors. This issue will ultimately make employee rating scales seem straightforward, but it is at present in the hands of physicists who have yet to regard this very complex judgement as a problem in itself.

The mismatch of the insensitive investigator and the elusive subject matter is, fortunately, self-correcting and limited to particular research problems. If the problem is important, progress will be made through the normal scientific processes of criticism and trial-and-error, although clear and simple answers will elude investigators simply because they do not exist.

This is currently a transitional period in which us 'hard-nosed' pioneers are mostly training students from softer disciplines. (I include myself because my arrival in 1961 almost qualifies, and my background in experimental psychology seems to furnish standards which generate a rather consistently negative history of my reviewing papers and grant applications.) In this transitional period, there is a widely held sense of failure in the field (at least, in the United States) which I read as a special symptom of the transition.

At a similar meeting several years ago, a senior statesman talked privately of how he felt he had oversold the field. From positions of influence he had promised that better techniques would move more and better information that would generate more and better science and technology. Analyzed from a rough epistemology, he, a 'hard science' type, dealt with science policy people, nearly all 'hard science' types. He expected to deliver, and they expected to receive, clear simple answers; he didn't and they didn't. Techniques, systems, hardware and software flourished, but the end product was disappointing.

From my perspective, this was not 'failure' but the painful recognition that information science, in looking for end results, is a 'soft' science because it studies processes and outcome of information use and analysis. In terms of examining the effects of information on people, their use and interaction with information in a wide variety of situations, the first, very approximate answers came in. However, the problems remained, and they are now beginning to look exceedingly difficult and complex.

A second related symptom of transition, at least in the United States, is a failure to develop much consensus on the quality of research. I won't dwell on this point but the difficulty seems to lie in hard types of scientists evaluating 'soft' types of persons proposing diverse solutions to the various unsolved structural, cognitive and language problems in the field.

To overview, information science is a 'soft' science. However, the motley crew who opened up and defined research problem areas are having difficulty recognizing that fact even though they are mostly in settings with 'soft'-er students and most research problems are now exhibiting a recalcitrant 'softness'. All of these factors mark this as a special period of transition with low scientific consensus and, at times, a sense of failure.

Let me add a further broad characterization of the field. Information science is and should be non-Paradigmatic, my own counter-Kuhnian term for fields that flourish on new facts and techniques rather than new conceptualization.[6] There is no sin in this, much of good science can be understood better as the unravelling or wresting of new facts from a recalcitrant nature than as a series of revolutions. Biochemistry hit upon the broad relationship between genetic information and biological function and structure a good time ago, probably prior to World War I, and this has set the stage for a continuing series of patiently achieved triumphs in research.[7] The only candidate for a revolution in the softer sciences is transformational grammar and now that is beginning to look like the re-evaluation of a European school of thought that had not originally been allowed to grow through adolescence. In the other social and behavioural sciences, there is a continuing intellectual dither but there has been little replacement of concepts or vocabulary for 30-40 years or longer.

Kuhn wrote an immensely valuable book for understanding certain transitions in physics.[8] It seems to turn into a potentially dangerous weapon, sending many of the untutored around looking for revolutions in softer areas - including ours. Most of the 'revolution seekers', as we might call them, are actually ignoring most of some fifty years of brilliant scholarship which has tried to comprehend why and how science is such a uniquely successful social endeavour. I feel we should require anyone who quotes or cites Kuhn, and not a single other writer in the history and philosophy of science, to read and critique all existing volumes of Needham's *Science and Civilisation in Ancient China* prior to allowing him further access to publication.[9]

Having said this much let me get myself into more trouble and try to delimit those subject matters and methods which should carry into the future. First, however, let me make clear that I am talking 'science', a coherent body of ideas and findings, not 'technology', their application and the

application of a variety of high skills, normally available for money.

First, information science has, as a science, only two sources of fundamental research problems, scientific and technical information and management information. All researchable questions - whether on organization, structure, language, cognition, etc. - seem to appear in their purest or most difficult forms in these areas. Each of these broad types of information feeds into systems that have output and feedback, features essential to assessing the quality of research results. Principles and techniques developed in these areas may be applied broadly, but they are never severely tested outside these areas.

In the first portion of this address, I made clear our inability to evade the human component; I have just indicated some of the special features of scientific and management information relative to the testing of research. There is a final critical feature, namely, that the field is as wedded to the large data processing devices and the large file as to the human being.

The evident progress and excitement surrounding statistical studies of language and documents is not an accident. A probabilistic approach and large numbers are inherent to the field - a matter which should not be entirely forgotten by either an investigator or a subspecialty. In many subspecialities of information science, the limited experiment should be pursued just far enough to make the point before going to a larger system or trial. Perhaps this apsect of information science explains the low impact of certain retrieval research. For example, through a series of 'invisible college' files and other groupings, I keep track of more documents in my office than, I believe, have been used in any 'test collection' for studying retrieval.

When we combine the machine, the file, and human being in a statistically based science we generate a series of problems dealing with structure, language and cognition that

relate to, but differ in several important respects, from similar problems in linguistics, psychology, sociology, and even epistemology. The disciplinary pressures in these other fields must force them to the simple problem, the 'pure' example, the uncontaminated result; pressures which must be counter-productive in our field. The final test of information science principles lies in large files, complex content and the powerful feedback mechanisms associated with scientific and management information.

In the above I have deliberately ignored the distinctive economic and organizational problems of information. While information scientists may work on these problems - and they are certainly important - the work will be congruent with and meet the standards and follow the directions of the economic and decision disciplines.

Let me approach the entire argument above - that following the discussion of 'softness' and my inability to anticipate an intellectual revolution - from a different point of view. One answer to the question 'What kind of science should information science be?' is that information science is simply the exploration, within certain limited parameters, of problems within operations research, within linguistics, within psychology and so on - and that one can adopt the internal standards of those disciplines. Instead, I say that the conditions of interest to us set special and more stringent standards - that, despite the similarity of subject matter, we should expect to establish a distinctive discipline with its own internal direction and standards. There are areas, of which I have noted two, and there may be more, in which our work is likely to retain the standard and direction of earlier established disciplines.

Having argued for the 'softness' and non-revolutionary character of information science, having discussed what it should do and roughly how, I would like to conclude with a discussion of my fears and hopes for the discipline.

First, fears: It would be too easy to answer the question "What kind of science should information science be?" with "It shouldn't be a science at all, it's a technology!" Your presence here, your patience with these remarks implies that we disagree with that idea. However, are we prepared to plan and work towards those intellectual and social conditions that mark mature sciences, even soft ones? At a minimum, they include the development of internal standards for scholarship and research and outside recognition of our ability to determine new directions of work. Both require high degrees of consensus among ourselves, but both are absolutely critical to achieve the degree of institutionalization required to sustain students and research.

Are we prepared to resist the direction set by society and the market place on occasion? If we are not, the field may become little more than a series of purchasable skills.

Ben-David and the emerging work on the history of biochemistry show these levels of institutionalization can be difficult to achieve, even when the fields are producing highly significant empirical findings.[10] Let me rephrase this - at many points in the history of science there has been little internal direction, little external support for the independence of such direction, and low institutionalization in terms of recognition as departments or schools, professorships and resources - even though the science produced is excellent. (Progress is made through an ancient and honoured scientific tradition dating to Galileo and before, now called bootlegging - getting monies for technology and doing science.) Does our work justify the institutionalization of information science? However we answer that question, looking into the past furnishes the outlines for a formidable program: Do good work, find and identify our excellences, establish standards, set directions, and convince others in our various national societies that we know our own business. This seems difficult but particularly necessary once we recognize the essential 'soft' and resistant character of our research problems.

Unfortunately, certain patterns of governmental support operate directly against this program and, instead, support only a diffuse technology. In discussing this, I will consider a parochial, ie American, example. I hope you will bear with me because I feel that this illustrates the way that well-meaning people can make virtually all the wrong decisions in terms of the ultimate development of the discipline.

A nice sampling of pernicious features have been embodied in recent programs of the Office of Science Information Service of the National Science Foundation of the United States. Some recent developments follow: After consultation with the information 'community', NSF invited proposals to complete prescribed tasks with mostly short-term support. In choosing the 'community' to be consulted there were no distinctions made between professionals and scientists. In a special project - conducted by the American Society of Information Science - local chapters were asked to meet and respond to the proposed program; the practices of the Chinese Cultural Revolution found an unexpected home, here.

Even if the 'community' had been scientists, communities don't do research; elites do research - a fact which we had all guessed even before Derek Price and the Coles made it crystal clear.[11] Second, determining the 'next steps' in science is the whole challenge of the game, and it is normally achieved by diversity of effort, trial-and-error, and exacting criticism. A government agency laying out tasks reveals an appalling level of intellectual arrogance. Finally, the odd and short time periods set to complete tasks eliminates most possibilities of supporting and training research students.

An additional note is that most of the prescribed tasks exist in that funny nether world of 'cost studies', 'demonstrations' and 'prototypes'. Neither tests of practicality nor exacting intellectual standards can be applied to such stuff, and I dread the outcome of generating a professional elite through their participating in these activities. My

dismay is further heightened by my beliefs, already developed in this paper, that most basic problems in the field are exceedingly difficult and will remain problems for years to come.

Let me summarize my fears for the field:

- continuing confusion between technology and science, professional and scientist

- most support involving external direction of activities

- misrepresentation of potential and nature of field in order to obtain support

- no support of research students

- support distributed over groups without long-term commitment to problem area.

When these patterns continue, the field cannot develop internal standards and direction or train new generations of students with any sense of 'craft'.

My hopes for the field are based, largely, on some changes consequent to changes in personnel and activities - in a sense, the outcome of normal aging processes for a discipline. Also, I am usually (sometimes blindly) optimistic.

Predicting the intellectual 'aging' of information science requires consideration of social conditions. The straightened conditions of higher education in many countries will not permit the burgeoning of any field. I anticipate some modest increases in the institutionalization of information science in terms of a few more departments and research groups, additional professorships and students. If I had my druthers, I would prefer to see information science attached to schools of librarianship (or documentation) and to business schools, preferably in diverse academic settings. Engineering schools and institutes typically, but not always, have rather circumscribed faculties, and, I am afraid, the

cognitive problems are likely to be ignored. Ideally, I would like people to have access to linguists, operations researchers, organizational theorists, mathematicians, historians and philosophers of science, and so on. It is the sort of thing that is easier done at Harvard Square in our Cambridge or at 30th Street and Market in Philadelphia than on the Front at Aberystwyth.

I believe there will be modest increases in funding fundamental researchers and research students. Other persons have noted the same difficulties I have and I understand that several European countries are trying to play 'catch up', which never really works but which does help the discipline. There will be a continuing softening of the personnel of the field and a similar modification of attitudes in assessing research. Some messy ideas will be given breaks that they may not deserve.

Turning now to the intellectual 'aging' of the field, conditions should favour several special and beneficial types of interactions among research problems, research skills, and people. The first type, scientific cross-fertilization, was earlier discussed by Derek Price in *Science since Babylon*[12] Price sees the whole of western science as an abnormal condition growing from the early successes of Ptolemaic astronomy. In turn, he sees this as a dramatic case of cross-fertilization in science, growing from strong, necessary, but utterly divergent roots in Greek geometry and Babylonian computational science. The reconstruction of many recent major discoveries, from DNA to superconductivity, seem to rely on almost chance encounters of different expertises. It may be merely a matter of the fellow in the office or lab next door looking in and seeing the problem with different eyes.

The persistence of problems in information science is producing diverse and interesting roamings and recombinations of people and skills. Thus, we have an astrophysicist, A.J. Meadows, dealing expertly with very soft issues in his

*Communication in Science*, and a wide range of persons moving into studies of the structure of scientific literature.[13] I anticipate that both substantial research achievements and better integration of the field will flow from such mobility.

There is a second type of crossfertilization which has even greater potential - because it creates new types of people, and which should have a special impact on information science - because of the diverse background of the 'pioneer' generation. We should make every effort to create meaningful and expert hybrids - for example, a student who works on the natural structure of scientific literature with me and on file structure with my colleague, Charlie Meadow, or upon classification theory and operations research. The idea here is, whether or not your student is better than you in your own field, make very sure that he is *much* better than you in another related field. This natural process is, for me, the chief hope of the discipline.

For completeness, I would like to mention another interaction of persons, skills and problems. We should begin, I believe, to consider the possible inter-relations of subspecialities. There is an implicit internal discussion in bio-medical science underlying a process by which problems are passed from one sub-speciality to another. Thus, a group of virologists may, both in print and informally, argue for the biological and medical importance of a particular material so as to persuade a biochemistry lab to work on it. Sometimes a biophysicist feeds a problem for another speciality, like crystallography. In these areas, the transfer is orderly and it is enforced by the nature of the materials and the limitations of the technique.

For information science, the question I wish to raise is: If we more strongly delineate subspecialities and stimulate an orderly transfer of problems, can we improve the discipline? I leave you with that question, reminding you that it argues against many 'systems approaches', which invite the professional to do a poor operations study, after he does a poor literature study, after he does a poor user study, etc.

Finally, I see one development of which I highly approve as an observer of sciences, and particularly, as an observer of our science. Many of the original pioneers are now clearly in the latter periods of their career. There is, in the United Kingdom, a good half dozen of these persons who perform the senior scientist role to perfection; they review, they criticize, they explicate - functions vital to sustaining a discipline. They give the rest of us the benefit of years of wrestling with data and issues. They are the envy of us in the United States who, I am afraid, wore out our equivalent generation flying them to and from Washington, to raise monies.

## FOOTNOTES

2. The primary sources of these concepts are:

Merton, Robert King: *The Sociology of Science: theoretical and empirical investigations*. Chicago: University of Chicago Press. (1973)

Price, Derek J. de Solla: "Citation Measures of Hard Science, Soft Science, Technology and Nonscience". In C.E. Nelson and D.K. Pollack; *Communication among Scientists and Engineers*. Lexington, Mass.: Heath. (1970)

Ben-David, Joseph: *The Scientist's Role in Society; a comparative study*. Englewood Cliffs, New Jersey: Prentice-Hall. (1971)

Ziman, John M: *Public Knowledge; the social dimension of science*. Cambridge, England: University Press. (1968)

Ravetz, Jerome R: *Scientific Knowledge and its Social Problems*. New York: Oxford University Press. (1971)

3. See, for example:

Forman, Paul: "Weimar Culture, Causality and Quantum Mechanics 1918-27: Adaptation by German Physicists

to a Hostile Intellectual Environment", *Historical Studies in the Physical Sciences, 3,* pp.1-115. (1971)

4. See Price *op cit* for discussion of this distinction.

5. Saracevic, Tefko: "On the Concept of Relevance in Information Science" PhD dissertation, School of Library Science, Case Western Rese University. (June 1970)

6. Kuhn, Thomas S. *The Structure of Scientific Revolutions.* 2nd ed. Chicago: University of Chicago Press. (1970)

7. The examples here and below, referring to biochemistry and biomedicine, are taken from Koyler, Robert,: "The history of biochemistry: a survey". *Journal of the History of Biology, 9,* in press 1975, and our own current work on contemporary science.

8. Kuhn *op cit*

9. Needham, Joseph: *Science and Civilization in Ancient China.* Vols. I-VI, Cambridge: Cambridge University Press.(1954)

10. Ben-David *op cit*

11. Cole, John and Cole, Stephen: *Social Stratification in Science.* Chicago: University of Chicago. (1973)

12. Price *ibid*

13. Meadows, A.J: *Communication in science.* London: Butterworths. (1974)

## ACKNOWLEDGEMENT

I wish to acknowledge the help of M. Carl Drott and Janet P. Griffith in discussing the issues and approach in the earlier drafts. With the first I have discussed the general issues for several years. The work was supported by Public Health Service Grant No. RO1 LM 00911-04.