



**QUEEN'S
UNIVERSITY
BELFAST**

Description and search labor for information retrieval

Warner, J. (2007). Description and search labor for information retrieval. *Journal of the American Society for Information Science and Technology*, 58(12), 1783-1790. <https://doi.org/10.1002/asi.20664>

Published in:

Journal of the American Society for Information Science and Technology

Document Version:

Peer reviewed version

Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

© 2007 Wiley Periodicals, Inc.

This is the peer reviewed version of this article, which has been published in final form at doi: 10.1002/asi.20664. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Open Access

This research has been made openly available by Queen's academics and its Open Research team. We would love to hear how access to this research benefits you. – Share your feedback with us: <http://go.qub.ac.uk/oa-feedback>

This is a postprint of an article published in the Journal of the American Society for Information Science and Technology copyright © 2007 (American Society for Information Science and Technology) at <http://onlinelibrary.wiley.com/doi/10.1002/asi.20664/abstract>

Description and search labor for information retrieval

Julian Warner

**School of Management and Economics
The Queen's University of Belfast**

j.warner@qub.ac.uk

Abstract. Selection power is taken as the fundamental value for information retrieval systems. Selection power is regarded as produced by selection labor, which can itself be decomposed into description and search labor. As forms of mental labor, description and search labor participate in the conditions for labor and for mental labor. Concepts and distinctions applicable to physical and mental labor are indicated, introducing the necessity of labor for survival, the idea of technology as a human construction, and the possibility of the transfer of human labor to technology. Distinctions specific to mental labor, particular between semantic and syntactic labor, are introduced. The high cost of human mental labor is also indicated.

Description labor is exemplified by cataloging, classification, and database description, can be more formally understood as the labor involved in the transformation of objects for description into searchable descriptions, and is also understood to include interpretation. Search labor is conceived as the labor expended in searching systems. For both description and search labor, there has been a progressive reduction in direct human labor, with its syntactic aspects transferred to technology, effectively compelled by the high relative costs of direct human labor compared to machine processes.

Introduction

Selection power, or the ability to make informed choices between objects or representations of objects, has been argued for as the primary aim for information retrieval systems, from a number of perspectives. Similar principles for retrieval have been developed in partly independent discourses, for instance in the value placed upon the discriminatory power of an index term in discussions of indexing and in the conception of bibliographic control as mastery over written and published records (Unesco/Library of Congress, 1950, p.1). Systems in common use embody facilities for enhancing selection power and the use of systems to this end, and the survival of such systems in the market for information services (Swanson, 1980, p.128), testifies to the perceived utility of selection power. Ordinary discourse comments from consumers of such systems also value control and selection. Understandings embodied in some significant scholarly discourses, in practice, and in ordinary discourse are, then, in advance of theoretical articulation, which continues to value query transformation above selection power, in the dominant research tradition. A link to deep, as well as to ordinary discourse or widely diffused, aspects of human experience is implied by the etymology of intelligence, from *inter-legere*, or to choose between (Stevens, 1998, p.66).

Selection power can be conceived as a fundamental concept, open to elucidation but not

to further decomposition into more primitive entities. It can also be understood as a quality of human consciousness, which can be assisted or frustrated by system design, but which does not inhere in the system itself. Under certain historical conditions, and levels of technological development, selection power is produced by activities, such as cataloguing, classification, description of objects for databases, and searching of catalogues and databases, which can all be comprehended and understood as selection labor. A fundamental proposition was developed, that selection power was produced by selection labor (Warner, 2007).

Selection labor was characterized as form of mental labor and theoretical minima established, for a given collection of objects, in a previous paper (Warner, 2007). The separation of selection labor into description and search labor with the *premodern* technologies of writing and printing on paper was noted. Similarly, the reconvergence of description and search activities with computer-based, or *modern*, technologies, together with the possibility of sustaining analytical distinctions between them, was acknowledged (Warner, 2007). The activities of description and searching still need to be more fully empirically characterized, as the components of selection labor.

The concern of this paper is, then, with developing and elucidating directly the concepts of description and search labor. Description and search labor, as mental labor, participate in the conditions for other forms of mental labor. Distinctions between types of mental labor and their different possibilities of transfer to technology need first to be introduced. Description and search labor, and their relation to selection labor, can then be understood, from the perspective on mental labor and technology established.

Concepts of mental labor

Labor and mental labor

Labor, or productive work in nature, is a fundamental condition of human life, imposed by the necessity for survival, both for *Genesis* and for Marx.

The labour process ... is purposeful activity aimed at the production of use-values. It is an appropriation of what exists in nature for the requirements of man. It is the universal condition for the metabolic interaction (*Stoffwechsel*) between man and nature, the everlasting nature imposed condition of human existence, and it is therefore independent of every form of that existence, or rather it is common to all forms of society in which human beings live.

(Marx, 1867/1976, p.290)

Mental labor has tended to be conceived as an adjunct to enhancing control over the physical environment, but can also be considered as an activity in itself, with the possibilities for its mechanization explored (Minsky, 1967, p.2).

Technology, including agrarian, industrial, and information technologies, can be regarded as a human construction, as the product of human labor on natural resources and preexisting humanly made products, rather than as naturally or objectively given. The understanding of technology as a human construction receives its fullest expression in Marx's work, in a classic passage from the *Grundrisse*, whose themes then implicitly inform the treatment of technology in the subsequent *Capital* (Marx, 1867/1976).

Nature builds no machines, no locomotives, railways, electric telegraphs, self-acting mules etc. These are products of human industry; natural material transformed into organs of the human will over nature, or of human participation in nature. They are *organs of the human brain, created by the human hand*; the power of knowledge, objectified. The development of fixed capital indicates to what degree general social knowledge has become a *direct force of production*.

(Marx, 1858/1973, p.706)

The focus of the passage is on industrial technologies, in accord with Marx's historical period, although automatic devices involving control mechanisms, 'self-acting mules', are mentioned. Regarding information technology as a human construction, concerned with the transformation of signals rather than natural resources (Warner, 2004, pp.5-35), indicates the possibility of a similar transfer of human labor to technology. Transfer of direct human labor to technology, for both industrial and information technology, can both speed processes and enable activities previously impossible.

The possibility of the transfer of direct human labor to technology enables a dynamic between human labor and its technological products, in which forms of direct human work are progressively transferred to technological processes. The dynamic is compelled by the historical search for greater control over the environment and is accelerated by the innovatory dynamic of capitalism, including the impulse from the reduced immediate costs of labor transferred to technology. From the perspective indicated here, the dynamic can be seen to apply to mental labor, as an activity in itself and not just as an adjunct to physical control over the physical environment. The social division of labor, for instance between master and slave or intellectual and clerical labor, can anticipate the division of labor between human labor and machine process.

Different types of labor can be distinguished, adopting a distinction made by Marx of universal from communal labor and applying it to informational labor, processes, and products.

We must distinguish here, incidentally, between universal labour and communal labour. They both play their part in the production process, and merge into one another, but they are each different as well. Universal labour is all scientific work, all discovery and invention. It is brought about partly by the cooperation of men now living, but partly also by building on earlier

work. Communal labour, however, simply involves the direct cooperation of individuals.

(Marx, 1894/1981, p.199)

As they merge into each other, universal and communal labor are not encountered in pure form, and can also be embodied in the products of labor, but their distribution can vary significantly. In this context, information technology in its machine aspect would be regarded as products primarily of universal labor, although constructed, activated, and renewed by communal labor; programs as products of communal labor in design and writing, developed from universal labor and its products, such as understandings of the algorithmic process and programming languages; and human description of information objects as primarily involving communal labor, guided by the universal labor embodied in codes for description. Communal labor has immediate costs, in human energy required, which may be matched by a wage, while the costs of universal labor have been absorbed historically. The distinction of communal from universal labor, as formulated by Marx, does involve mental labor – ‘Universal labour is all scientific work, all discovery and invention’ (Marx, 1894/1981, p.199) – and can also be applied directly to mental labor and to its products, in their modern manifestations.

Labor, process, and product can be explicitly distinguished, with process and product separating out from an originally undifferentiated labor (in unrecorded oral speech, for instance labor and process are not distinguished and the product disappears with its creation). Once the categories are separated out, labor can be understood as direct human work and process as the activity abstracted out from labor, with the possibility of transfer to technology. An originally partly nominal distinction of labor from process, arising from unease at applying the term *labor* to inorganic activities, also has substantive implications, with the possibility of greater rigidity and exactness in the process compared to direct human labor, partly arising from the preliminary formalization. Products, in this context semiotic products, such as catalog records, are products of direct human labor and of technological processes.

Some supporting assumptions for the argument here, derived from the consideration of human labor and its relation to technology, itself conceived as a human construction, can, then, be summarized. Human labor, including human mental labor, can be transferred to technology. The distribution between human work and technological process varies historically, as communal labor is transferred to products derived from universal and communal labor. The most significant possibility, for information retrieval, is that human mental labor can be transferred to technology.

Distinctions specific to mental labor, not captured by the historical concern with physical and productive labor, can now be introduced.

Distinctions within mental labor

The social division of intellectual labor is more familiar than a separation between aspects of mental tasks, first as a division between the priesthood and those engaged in more directly productive labor (Childe, 1936), and developing with the differentiation of disciplines and discourse communities (Goody and Watt, 1968).

The 19th century witnessed the deliberate division, in real world practice, of aspects of mental tasks between people working within organizations or on extensive projects, in addition to the established social division of mental labor. A theoretical understanding of the possibilities of the division and mechanization of mental labor also began to emerge. At a level mediating between practice and theory, the production of tables of logarithms was found to be amenable to manufacture by the subdivision of the tasks involved and the division of the human labor for those tasks. Prony, encountering Adam Smith's work on the division of manufacturing labor, by chance in a bookshop, conceived of this possibility:

par une espèce d'inspiration, il conçut l'expédient de mettre ses logarithmes en *manufacture* comme les épingles.

(Babbage, 1835/1963, p.93)

An interesting, and potentially disturbing, feature discovered in the actual process of production of the tables was the relative accuracy of clerical and intellectual human labor.

Persons [nine tenths with no knowledge or arithmetic beyond addition and subtraction] were usually found more correct in their calculations, than those who possessed a more extensive knowledge of the subject.

(Babbage, 1835/1963, p.195).

Considerations of meaning can, then, for human computers, distract attention from simple computational operations. At a more deliberately theoretical level, Babbage himself, similarly influenced by and adapting Adam Smith on the division of labor in manufacturing, noted the possibility for the division of mental labor.

The division of labour is no less applicable to mental productions than to those in which material bodies are concerned.

(Babbage, 1835/1963, p.379)

Mental labor has a material aspect, particularly in the use of exosomatic technologies, and it is this material aspect which gives the possibility of mechanization.

Mechanical mental labor is contrasted with mental labor belonging to the 'the domain of the understanding, requiring the intervention of reasoning' (Babbage, 1989, p.246),

closely involving considerations of meaning. For Babbage, one impulse to the design and construction of computing machines was to transfer the mechanical part of the mathematician's labor to automatic machinery (Babbage, 1989, pp.246-247). A similar impulse can also be discerned in the diffusion by adoption of modern information technologies, a process also influenced by a desire to avoid the costs of direct human labor, or communal labor in Marx's terms.

In a development from the familiar, and deeply culturally embedded, distinction of semantics from syntax, semantic and syntactic mental labor have been differentiated. Semantic labor is concerned with transformations motivated by the meaning or signified of symbols, while syntactic labor is determined by the form alone of symbols, operating on them in their aspect as signals. Semantic labor requires direct human involvement while originally human syntactic labor can be transferred to information technology, where it becomes a machine process. Direct human labor has high costs while mental labor transferred to technology is likely to have relatively diminished costs, under modern conditions. The distinction of semantic from syntactic labor has an analogue in ordinary discourse, and the transfer of syntactic mental labor to technology occurs in everyday practice, suggestive of the robustness and wide applicability of the distinction, despite its only recent theoretical articulation (Warner, 2005).

The distinction between semantic and syntactic mental labor can be conveyed here by an example from a literary source, occurring chronologically close to and within the same broad cultural context as Boole's formalization of logic, conceived as the laws of thought (Boole, 1854).

'My other piece of advice, Copperfield,' said Mr. Micawber, 'you know. Annual income twenty pounds, annual expenditure nineteen nineteen and six, result happiness. Annual income twenty pounds, annual expenditure twenty pounds ought and six, result misery. The blossom is blighted, the leaf is withered, the god of day goes down upon the dreary scene, and – and in short you are for ever floored. As I am!'

(Dickens, 1850/1966)

In this passage, the syntactic process of calculation is parodied by the substitution of a semantically congruent, although syntactically dissonant, result (a curious, and revealing inversion, of modern experience with computer operations on written text for information retrieval and spell-checking). In contrast to the syntactically generated result it implicitly and replaces, the semantic result is not generalizable to other similar syntactic procedures. In 19th century practice, human mental labor would have been assisted by the exosomatic technologies of writing and both semantic and syntactic labor would have involved continuous human intervention. Following the late 20th century mechanization of mental labor, syntactic labor can be transferred to information technology, operating deterministically between intervals of human intervention, opening up and revealing a distinction between semantic and syntactic mental labor. Attempts to transfer human mental labor to information technology, with the exception of processes which were

already formalized or known to be formalizable, have tended to reveal the complexity and intractability of semantic labor.

Support for the distinction of semantic from syntactic labor, and of labor from process, can be derived from the implicit endorsement of strongly analogous distinctions by the United States Supreme Court, in a highly significant copyright judgment involving modern information technologies. In *Feist vs. Rural* (Feist, 1991), copyright protection was denied to telephone white pages.

Nor can Rural claim originality in its coordination and arrangement of facts. The white pages do nothing more than list Rural's subscribers in alphabetical order. This arrangement may, technically speaking, owe its origin to Rural; no one disputes that Rural undertook the task of alphabetizing the names itself. But there is nothing remotely creative about arranging names alphabetically in a white pages directory. It is an age-old practice, firmly rooted in tradition and so commonplace that it has come to be expected as a matter of course. ... It is not only unoriginal, it is practically inevitable.

(Feist, 1991)

The 'age-old practice' corresponds to syntactic labor and the alphabetization of names to a syntactic process, delegated to the forms of information technology current in the 1980s. Slightly obscured within the judgment, and not fully brought into contrast with the explicit reversal of the labor theory of copyright is a reference to 'writings which are to be protected ... [as] the fruits of intellectual labor, embodied in the form of books, prints, engravings, and the like' (Feist, 1991). The intellectual labor embodied in the protected writings is closely analogous to semantic labor. The judgment has independent interest, but, in this context, its immediate value is as a wider public analogue to the distinctions of syntactic and semantic mental labor, and of human labor from machine process, supporting their validity.

The understanding of the conditions for mental labor can, then, be summarized. Human labor, including mental labor, can be transferred to technology, undergoing a transformation into a machine process. Human mental labor can be semantic or syntactic in character, directly motivated by considerations of meaning or reduced to pattern governed transformations. Only syntactic, and not semantic, labor can be transferred to information technology. Description and search labor, as forms of mental labor, participate in the conditions for mental labor.

Description and search labor

Selection labor

The historical separation out of description and search labor from selection labor, and their current reconvergence, supported the proposition that, selection labor can be significantly distributed between description and search labor but that its overall sum

could not be diminished below certain theoretically established limits (Warner, 2007). The understanding of description and search labor as forms of mental labor supports the further proposition: that the syntactic aspects of description and searching can be transferred to technology but not their semantic components. Units of syntactic labor, if they were distinguished, would have some analogy with the bit, or binary choice, and can be realized as direct human syntactic labor or transferred to a technological process.

Description labor

Description labor is, then, understood as one component of selection labor. As mental labor it can have a material and exosomatic aspect and semantic and syntactic components. Description labor will be considered in a more directly empirical and contemporary fashion than selection power and selection labor (Warner, 2007), but the historical emergence of description labor in information systems with the emergence of written literacy is acknowledged. Equally, the possibility of its reduction, its transfer to technology as process and its absorption within selection labor, with selection labor becoming a more substantive category, is acknowledged. The labor embodied in documents described is largely treated as a given, rather than fully explored, but the contrast between the technology of writing on paper, demanding separate description, and of computation, enabling automatic generation of syntactic descriptions, is fully incorporated into the schema developed. Human description labor is understood to include interpretation as well as apparently more simple forms of description, classically understood to be embodied in the activities of classification and of cataloguing.

Description labor can be understood, first ostensibly and empirically, as the work involved in such processes as cataloging, classification, indexing, and database description. More analytically, although still consistently, it can be conceived as the work involved in the transformation of objects for description, which could be documents, images, or people, into searchable descriptions, for the purposes of subsequent retrieval. Two aspects of the transformation of objects for description can be distinguished: the description of objects and the assembling of these descriptions into searchable lists or indexes. These aspects may merge into each other in practice but can be analytically separated. The implicit aim of description labor is to increase selection power and a further effect of increasing selection power can be to reduce labor expended in searching.

The origins of the separation of syntax from semantics, with written language (Warner, 2005, pp.557-563), may have continuing implications for the relative effectiveness of applying syntactic, or pattern-transforming, procedures, to verbal and non-verbal graphic signs. Non-verbal signs can be subjected to syntactic or pattern-based transformations, with those transformations realized as labor or as process, but it has been difficult to endow those transformations with semantic significance either for similarity and difference between signs or for establishing orders for display. Technically, it would not be difficult to transform a digital photograph into a searchable description (distribution of colors in the image for instance) and to establish measures of similarity between images. It has proved far more difficult to create meaningful searchable descriptions or measures

of similarity, except with sets of images produced under highly controlled conditions and thereby effectively selected for their formal similarity to one another. Images, including photographs, received as iconic signs, with a motivated resemblance from signified to signifier, could lend themselves to iconic representation for display and retrieval, possibly reducing the size of images in order to enable scanning. Indexing and retrieval for *Google Advanced Image Search* is based primarily on the verbal descriptions of the images, not on matching between the graphic images, and iconic modes of representation are used for the display of retrieved results. In contrast, written verbal signs which embody the distinction between semantics and syntax are more amenable to syntactic operations, which can incorporate, in partial and incomplete ways, their semantic significance.

Examples of significant information systems can reveal the reduction in direct human labor and the increase in the fullness of descriptions. In late 19th and early 20th century practice, *Palmers Index to The Times* was created by direct human labor, assisted by current technologies, particularly writing and printing. For instance, the headline, ‘Mad dogs in Sheffield’, would be transformed into the index entry, *Mad dogs in Sheffield*, and filed in a list of entries for that year. The process of description is syntactic, although conducted by clerical labor rather than by machine, with little semantic intervention. The storage constraints of the technologies compel the production of descriptions more concise than the objects described. Technologies also have to be used in a primarily non-deterministic mode, with continuous human intervention, particularly for the construction rather than printing of indexes. Other late 19th and early 20th century information systems, most obviously the construction of the *British Museum Catalogue*, involved substantial human semantic intervention in the making and listing of descriptions, with, for instance, reduction of names to canonical form and structured grouping of material by and about significant authors, although the subject approach was not fully endorsed (Roberts, 1977).

These contrasting descriptive practices have modern descendants, which can be embodied within the same information system. *Google Advanced Search* could be regarded as the descendant to the practice of indexing newspapers, by syntactic transformations on the object-language. Further semantic work is not applied to the objects described, although these may contain deliberate descriptive elements, for instance, in the application of metadata. Descriptions, which may not be available for direct inspection, are automatically generated from the verbal objects themselves, with the further generation of indexes to those descriptions. The direct human clerical labor involved in 19th century practice has been delegated to modern information technologies, operating syntactically and deterministically. The need for more concise descriptions, imposed by storage constraints, has also been removed. The searcher’s labor remains intense.

WorldCat embodies both practices. Descriptions are created by human semantic work, guided by codes developed from their 19th century antecedents, on the verbal objects described, with the aim of creating more systematic descriptions than those which would be given by transcription of the verbal objects. These systematic descriptions aim to increase the selection power of the searcher and reduce their labor. Elements of syntactic

labor, transformed into a machine process, can also be found. For instance, humanly created descriptions would be automatically transformed into searchable indexes. Increasingly, descriptions themselves are also incorporating information derived by syntactic processes from the verbal objects taken for description.

Some common trends over time can be discerned in these processes of description. There is a decrease in the direct human labor involved in the description of objects, and, more intensely, in the compilation of indexes from these descriptions. In particular, syntactic components of description are being transferred from human labor to machine process. Descriptions are becoming more full and exact, enabled, although not compelled, by the reduction in technologically imposed storage constraints. Fullness and exactness can, analytically and substantively, be distinguished from discriminatory power and informativeness and may not enhance selection power, although they may aid specificity in searching. The social division of mental labor, between clerical and intellectual roles, has been progressively transformed into the division of labor between human work and machine processes, with syntactic labor delegated to technology. There are also indications of a transfer of labor from the producer to the searcher. Less human labor is being expended at the point of production of the system, in the description of objects for retrieval, and more human labor may be expended in searching. Contrasts between late 19th century and early 21st century practice arise from the development and adoption of information technologies, the progressive transformation of communal labor into the products of primarily universal labor (Warner, 2005, p.477), while continuities follow from the predominantly symbolic nature of the signs described.

A particular feature, combining continuity and change, is the transformation of the social division of labor into a division of labor between human and machine, with clerical or syntactic labor transferred to machine process. This pattern might have implications for the nature of expertise in searching information systems. At one extreme, understanding of the forces determining the construction of systems may not correlate with effectiveness in searching systems¹. Particularly acutely relevant is the changed relation between the language of discourse and the language of representation between premodern and modern information systems: a semantically assigned language of representation need not be the entry vocabulary, although later recourse could be made to it for its generic scope. The *prima facie* case would be that expertise in the language of discourse is likely to be stronger among domain than information specialists.

The costs of the human description labor involved in creating records for catalogs have been considered, although not fully within the conceptual framework developed here, and these considerations can provide empirical data to inform the argument. The costs of creating a catalogue record to the standards required for *WorldCat* have been estimated to be in the order of US \$40 (Hayes, 2000). From the perspective here, this cost can be regarded as the cost of human semantic description labor and the use of *WorldCat* records by participating libraries as the distribution of the products of labor, in order to share the costs of that labor. The books of OCLC show no value for *WorldCat* or for the costs involved in the creation of records directly by OCLC (Hayes, 2000, p.76). The failure to recognize the cost of the direct and accumulated intellectual labor embodied in catalog

records and catalogs has been regarded as ‘serious mis-measurement of phenomena of fundamental economic importance’ (Hayes, 2000, p.73).

The understanding developed of description labor can now be summarized. Description labor can be semantic or syntactic in character. Semantic description labor is exclusively directly human, in accord with the understanding developed of human mental labor and the possibilities of its transfer to technology. Syntactic description labor, in contrast, can be delegated to technology, where it is transformed from organically conducted labor into a machine process.

Search labor

Search labor was contradistinguished from description labor, with both description and search labor regarded as components of selection labor. Search labor can also be considered as a form of mental labor, with material, and semantic and syntactic, aspects. Within the schema developed, selection labor can be significantly distributed between description and search labor and its syntactic aspects transferred to technology, but its overall quantity cannot be diminished. With premodern systems, search labor might have emerged predominantly as search expertise, reflecting the substantial investment of human semantic labor in description processes. A signal and revealing, exception to the would be the arduous work, coupled with accumulating expertise, required for the searching for the sources for the construction of a subject bibliography, attempting exhaustivity and encountering bibliographic scatter (Greg, 1959; Bradford, 1948/1971) ². For analytical clarity, and with substantive justification, search labor can be regarded as the reverse or obverse, or even mirror image, of description labor.

The material aspect of search labor would be reflected in the psychomotor skills required for the operation of the keyboard, mouse, or other interface devices. The apparent complexity of this aspect of searching has been diminished by the progressive naturalization of information technologies and by improvement of the interface by system design. Both perceived complexity and complexity more inherent in the process have diminished, although human desires may also coevolve with technological provision. The recent relative stability of interface technologies may indicate that a teleological state has been reached, at least temporarily, after the late 20th century revolution in the mechanization of mental labor.

Syntactic processes themselves can be distinguished from syntactic aspects of searching. Syntactic processes would be represented by such activities as ordering of retrieved records or elimination of duplicate records and have been increasingly transferred to technology. With premodern technologies, such processes, for instance ordering of records, by date or by author, would have been accomplished by description labor or by direct human syntactic labor at the point of retrieval. For modern technologies, different orderings can be invoked, with labor transferred to machine processes, in searching.

Syntactic aspects of searching, in contrast to syntactic processes, would correspond to understanding of syntax and this could itself be considered a form of semantic

understanding. Specific components for understanding would include Boolean logical combinations, and their computational derivatives, their realization in system commands, and the likely effects of specific combinations. Boolean logic has been considered to be difficult to grasp, although its difficulty may have been exaggerated³. With the adoption of computational technologies, some understanding of Boolean logic has also diffused into public consciousness. Boolean operators are characteristically realized in contrasting system commands, although the underlying commonality should be recalled. Early studies indicated that the number of system commands used had little effect on system performance (Barraclough, 1977), although differences in commands were known to make it difficult for searchers to adapt from one system to another. In common with the material aspects of searching, syntactic aspects have been amenable to amelioration, both through system design and coevolution of consciousness.

Semantic components would centre on the translation of a topic into a searchable query and this has long been acknowledged as the most complex aspect of retrieval (Roberts, 1977; 1989). The process may remain difficult, but its nature, the social distribution of expertise, and distribution of labor between description and searching may be changing, with alterations in description processes. Semantic understanding and expertise with premodern technologies would involve understanding of the type and particular characteristics of the description labor applied to those documents, including the language of subject representation. With modern systems, understanding of the language of discourse of the documents represented, and the effects of generating automatic representations of those documents, may become more significant, with an associated social redistribution of expertise towards those fully familiar with the language of discourse. For some systems, for instance, a directory, there is increasingly less system imposed requirement to articulate a topic fully and verbally in advance of searching.

A further level of semantic understanding would be knowledge of the overall system of information system production, and the role or position of particular information systems within the used in that system (Roberts, 1989) (the term *system* should be taken to refer to an interacting set of components rather than a deliberately coordinated or planned system (Roberts, 1977)). Both the overall system, and the nature of expertise required, are changing with the diffusion of Internet search engines.

Search labor can, then, be understood as parallel to description labor. Search labor can be either semantic or syntactic in character. Semantic search labor is inescapably directly human, while syntactic search labor can be transferred to technology, where it becomes a machine process. The parallelism of search to description labor, which has been obtained without a Procrustean fitting of data to theory and without disguising their contrasts, reinforces the analogies between them. It also strengthens the conception of the possibility of the interchange of human labor and machine process between description and search, within the encompassing activity of selection labor.

Description and search labor

The dynamic which compels the transfer of syntactic labor to technology is connected with the costs of direct human labor and was acutely noted by Norbert Wiener, in a proleptic remark.

the automatic machine ... is the precise economic equivalent of slave labor. Any labor which competes with slave labor must accept the economic conditions of slave labor.

(Wiener, 1954, p.162)

With the reduction in semantic description labor, semantic labor may be transferred from description to searching and search labor increased. The process of searching may remain permanently intractable, at the semantic level.

Conclusion

The concepts of description and search labor have been elucidated, with parallel distinctions made for the work involved in description and searching. The possibility of the transfer of human mental labor to technology was first established. Semantic and syntactic mental labor were distinguished, with syntactic labor transferable to technology, as machine process, while semantic labor remained irreducibly human. Description labor was exemplified as cataloging, classification, and database description, with its semantic and syntactic aspects distinguished. Search labor was considered as strongly parallel to description labor. Selection labor, or the total labor involved in information retrieval, was amenable to significant distribution between description and searching, and its syntactic aspects could be transferred to technology, but could not otherwise be diminished.

A further stage in the development of a labor theoretic approach to information retrieval would be to review and synthesize all the elements of the argument, from selection power to selection, description, and search labor, and fully to introduce the dynamic compelling the transfer of human syntactic labor to technology, stemming from the costs of direct human labor. The parsimony and power of the overall argument and its correspondence to real world decision practice would also be revealed.

References

- Babbage, C. (1835/1963). On the economy of machinery and manufactures. 4th edition. London: Charles Knight, 1835. In Reprints of Economic Classics. A.M. Kelley: New York.
- Babbage, C. (1989). Science and reform: selected works of Charles Babbage. Edited with an introduction by A. Hyman. Cambridge etc.: Cambridge University Press.
- Barraclough, E.D. (1977). Online searching in information retrieval. *Journal of Documentation*. 33, 220-238.
- Boole, G. (1854). An investigation of the laws of thought, on which are founded the mathematical theories of logic and probabilities. London: Walton and Maberly and Cambridge: Macmillan.
- Bradford, S.C. (1948/1971). *Documentation*. First published 1948. 2nd edition. High Wycombe, Bucks., University Microfilms for the College of Librarianship Wales.
- Childe, V.G. (1956). *Society and knowledge*. London: George Allen & Unwin.
- Dickens, C. (1850/1966). *The personal history of David Copperfield*. Edited by J. Tambling. London: Penguin.
- Feist. (1991). U.S. Supreme Court. *Feist Publications, Inc. v. Rural Tel. Service Co.*, 499 U.S. 340 (1991). 499 U.S. 340 *Feist Publications, Inc. v. Rural Telephone Service Co., Inc. Certiorari to the United States Court of Appeals for the Tenth Circuit no. 89-1909*. Argued January 9, 1991. Decided March 27, 1991.
- Goody, J. and Watt, I. (1968). The consequences of literacy. In J. Goody editor. *Literacy in traditional societies*. Cambridge: Cambridge University Press.
- Greg, W.W. (1959). *A bibliography of the English printed drama to the restoration*. Volume IV. Introduction, additions, corrections, index of titles. London: printed for the Bibliographical Society at the University Press, Oxford.
- Hamer, R. (1970). *A choice of Anglo-Saxon verse*. London: Faber and Faber.
- Hayes, R.M. (2000). Assessing the value of a database company. In B. Cronin and H.B. Atkins (Eds.), *The web of knowledge: a festschrift in honor of Eugene Garfield* (pp.73-84). Medford, NJ, 2000: Information Today, Inc..
- Marx, K. (1858/1973). *Grundrisse: foundations of the critique of political economy* (Rough draft). Translated with a Foreword by Martin Nicolaus. London etc.: Penguin Books in association with New Left Review.

Marx, K. (1867/1976). *Capital: a critique of political economy*. Volume One. Introduced by E. Mandel. Translated by B. Fowkes. Harmondsworth etc.: Penguin Books in association with New Left Review.

Marx, K. (1894/1981). *Capital: a critique of political economy*. Volume Three. Introduced by E. Mandel. Translated by D. Fernbach. Harmondsworth etc.: Penguin Books in association with New Left Review.

Minsky, M. (1967). *Computation: finite and infinite machines*. Englewood Cliffs, NJ: Prentice-Hall.

Roberts, N. (1977). *Communication and the bibliographical system of the social sciences*. In N. Roberts (Ed.). *Use of social sciences literature* (pp.1-27). London: Butterworths.

Roberts, N. (1989). Online: in an educational cul-de-sac? *Education for information*. 7, 101-106.

Stevens, A. (1998). *Ariadne's clue: a guide to the symbols of mankind*. London etc.: Allen Lane, The Penguin Press.

Swanson, D.R. (1980). *Libraries and the growth of knowledge*. In D.R. Swanson (Ed.). *The role of libraries in the growth of knowledge* (pp.112-136). University of Chicago Press.

Unesco/Library of Congress. (1950). *Bibliographical services, their present state and possibilities of improvement: report prepared as working paper for an international conference on bibliography*. Washington, DC.

Warner, J. (2004). *Humanizing information technology*. Lanham, MD: Scarecrow Press.

Warner, J. (2005). Labor in information systems. *Annual Review of Information Science and Technology*. 39, 551-573.

Warner, J. (2007). Selection power and selection labor for information retrieval. *Journal of the American Society for Information Science and Technology*. To appear 2007.

Wiener, N. (1954). *The human use of human beings: cybernetics and society*. Revised edition. New York: De Capo Press.

Notes

1. I am indebted to a student for this observation.
2. W.W. Greg's comments after completing, *A Bibliography of the English Printed Drama to the Restoration*, the 'product of a lifetime of study', are worth recalling, with regard to the prolonged labor of searching, the accumulation of expertise, the subjectively experienced limitations of that expertise in fully comprehending the topic sought, and the modification of originally held intentions through interaction with the material discovered.

lest anyone should think that looking back on my work I feel any complacency over the manner of its execution, I here admit that I can hear the caustic critic who ever sits like a familiar imp at my elbow maintaining that my problem in writing this introduction has been threefold: first to discover what in fact I have done, next why I did it, and lastly how best it may be defended.

The decision to limit the work to printed plays had been one 'convenience or expediency' (Greg, 1959, p.v).

3. Some riddles informally anticipate the Boolean operators, particularly And (consider the Anglo-Saxon examples in Hamer, 1970, pp.95-107), but also Or and Not.

Acknowledgments

I would like to acknowledge the John Campbell Trust who supported a presentation of the themes of this article at the Annual Meeting of the American Society for Information Science and Technology Annual Meeting, Providence, Rhode Island, November 2004 and the hospitality and assistance of the Research Centre for the Social Sciences (RCSS), University of Edinburgh, where the article was originally written during study leave from the Queen's University of Belfast, February–July 2005.