2-2022

Context, Relevance, and Labor

Wayne de Fremery
Dominican University of California, wayne.defremery@dominican.edu

Michael K. Buckland
University of California, Berkeley

https://doi.org/10.1002/asi.24631

Received: 30 June 2021 | Revised: 2 January 2022 | Accepted: 10 February 2022

DOI: 10.1002/asi.24631

RESEARCH ARTICLE

JASIST WILEY

Context, relevance, and labor

Wayne de Fremery[1,2] | **Michael K. Buckland**[3]

[1]School of Media, Arts, and Science, Sogang University, Seoul, South Korea

[2]Barowsky School of Business, Dominican University of California, San Rafael, California, USA

[3]School of Information, University of California, Berkeley, California, USA

Correspondence
Wayne de Fremery, Barowsky School of Business, Dominican University of California, 50 Acacia Ave., San Rafael, California, USA
Email: wdefremery@gmail.com

Abstract

Since information science concerns the transmission of records, it concerns context. The transmission of documents ensures their arrival in new contexts. Documents and their copies are spread across times and places. The amount of labor required to discover and retrieve relevant documents is also formulated by context. Thus, any serious consideration of communication and of information technologies quickly leads to a concern with context, relevance, and labor. Information scientists have developed many theories of context, relevance, and labor but not a framework for organizing them and describing their relationship with one another. We propose the words *context* and *relevance* can be used to articulate a useful framework for considering the diversity of approaches to context and relevance in information science, as well as their relations with each other and with labor.

1 | INTRODUCTION

Context, relevance and labor are central concerns in information science. The purpose of this essay is to explore how they can be related to each other. Simply stated, relevance depends on meaning, meaning depends on context, and the labor required for search and access is closely tied to context. Although we are not primarily concerned with explaining the history of the of these three concepts in the field, some discussion of how the terms have been used is necessary to clarify relationships. In particular, the unsettled understanding of context requires some attention. Moreover, our goal of articulating a theoretical framework for organizing the diversity of empirical work done on context, relevance, and labor requires a more abstract level of discussion than is usual in information science.

If information science is concerned with "the transmission of the universe of human knowledge in recorded form" (Saracevic, 2017, p. 2216), then it is concerned with context since transmitting records ensures they will arrive in a new context. Information technologies, notably writing, printing, telecommunications, and the many forms of new media, are designed to formulate documents in distinct places and times (de Fremery & Buckland, 2022). As Shannon (1948, p. 379) wrote, the fundamental problem of communication concerns reproducing at one point "either exactly or approximately a message selected at another point." Shannon famously excluded a discussion of meaning from his theory to focus instead on a theoretical solution to the engineering problem of making a message produced at one point available as copy at "another point" and, thus, in a new context. Although he did not exclude its discussion explicitly, as with meaning, Shannon did not concern himself with the nature of what might contextualize a finite set of messages that could be selected at a source or the alternate context provided by a destination. Nor did he did concern himself with the relevance of a message at its source, its destination, or how its relevance is likely to be distinct at these two points. But just because Shannon did not deal with them explicitly in his essay, does not mean that meaning, relevance, and context are

inconsequential when considering messages. Concern with context arises throughout information science from indexing (e.g., keyword in context) through library classifications and information behavior to domain analysis.

Context and relevance have, of course, been central concerns of information science. Like Shannon's theory of communication, although not often described in such terms, theories of context and relevance have been formulated by the idea of documents separated by time or space. Where Shannon was concerned with the problem of reproducing a message, theorists of context and relevance have focused on the problem of associating messages (queries and records, e.g.) of different origins and separated by space or time so that they can be available in a single location at a moment in time. The amount of labor required to discover and retrieve messages, documents, and records produced at alternate places and times has been a significant concern. Thus, any serious consideration of communication and of information technologies leads quickly to a concern with contexts, with messages and documents produced or selected in one place or time and made available at another, then with relevance (judgments about messages or parts of messages that are or may be useful) and labor (the work of discovering and making messages or their parts accessible).

Context has long been an important topic in information science, and it has been approached in a variety of ways: as objective accounts of the circumstances of some event, as subjective perceptions of those same circumstances, or as the states-of-mind of human actors. As we discuss below, Agarwal (2018) provides a convenient and detailed overview of some of the many ways that concepts of context have been developed in relation to specific modes of investigation within information science. But these theories have not coalesced into a framework that can account for the variety of ways that context is approached. Four recent and significant information-related encyclopedias do not contain a definitional or even a historical article describing context and the diversity of ways it has been approached. (See the *Encyclopedia of Library and Information sciences* [McDonald & Levine-Clark, 2018], the *Handbook of Information Science* [Stock & Stock, 2013], *Information: A Historical Companion* [Blair et al., 2021], and the *ISKO Encyclopedia of Knowledge Organization* [Hjørland, 2021].) This might be contrasted with the situation in adjacent disciplines, such as pragmatics, where handbooks have comprehensive accounts of context (see Fetzer, 2017). The related notion of relevance, which has also been articulated in relation to specific information science problems, especially information retrieval, and has also resisted a satisfactory shared framework (White, 2018).

The limited progress toward a framework that helps to organize discussions of specific kinds of context is, we suggest, associated with a tendency to treat context empirically as an a priori state or situation. Even when contexts are understood to be dynamic, instances of context are selected to stand for a generalizable theory of context when, in fact, they are formulated in relation to specific information problems. When investigated empirically, contexts are found to have many elements and to be multidimensional (Agarwal, 2018, p. 98). Something more abstract is needed to develop a clearer general framework for organizing context as a useful concept in information science, to understand why theorizing context has been difficult, and to theorize the relationship between context and closely related concepts such as relevance and labor.

The challenge is that a focus on context, on what is connected at the periphery of an initial focus, reveals new contexts. A word in the middle of a book is read in the context of the those before it and after it. Those words are read in the context of the sentence in which they appear. The sentence is read in the context of the sentences before and after it. The sentences are read in the context of the book. The book can be investigated in the context of those beside it on the self or the language in which it is written or the author who wrote it or the publisher produced it, the book seller who sold it, and the other people who read it. This is to acknowledge that any discussion of context will be incomplete. It is also one reason why some early models for digital preservation have called references to information outside of their control, but upon which they depended, "Gödel ends" after the logician best known for his incompleteness theorems, Kurt Gödel (Cedars Project Team, 2001). Digital preservation, it turns out, like every topic in information science, concerns context. Just as Kurt Gödel arrived at a theoretically useful description of mathematics by using numbers to enumerate numbers and mathematical operations, we suggest that the word *context* itself, its etymologies and uses, provides a theoretically helpful way to organize concepts of context in the context of information science. The word is useful for thinking about how contexts are individuated in distinct branches of information science as well as the infinity of contexts inhabited by the people and phenomena that information science studies. The word *context* in this situation is also useful because it suggests a means for connecting the variety of ways that context has been considered so that a more, if never entirely complete framework for considering context in information science can be conceived. Rather than assuming contexts exist a priori and attempting empirical analysis, we adopt a more abstract approach. Doing so, we can recognize that what counts as relevant context for digital preservationists is distinct and may only be loosely connected to the contexts considered relevant in information retrieval or the contexts considered by those at work on information theory and the engineering problems of communication. At the same time, the connection enables the possibility that what counts as

context in information retrieval can inform those working on preservation and information theory and vis-versa. This is to say what is obvious but hard to see, that the word *context* can help to organize discourses about context because, as an abstraction, it can be inclusive of contexts' empirical and experienced multitudes.

Information science is centrally concerned with documents and copies (de Fremery and Buckland, 2022). Our discussion will focus on how they shape context and are shaped by context, the ways they become relevant, and how they affect labor. We use "document" in a broad sense as any object regarded as actually or potentially signifying something (Buckland, 2017). The quality of being a document—documentality—makes documents a form of evidence (Buckland, 2014). It is a characteristic of evidence that it points to something beyond itself, as when smoke indicates fire (Kelly, 2016). Indeed, social psychologist Robert Pagès (2021) insists that documents are symbols and philosopher Hans-Georg Gadamer suggests that symbols are documents. The "meaning of the symbolon" he writes, "depends on its physical presence and acquires a representational function only by being shown or spoken" (Gadamer, 1985, p. 65). Documents, therefore, can be more than mere representations. They can be performative in the sense of enabling, in the right circumstances, certain outcomes through their presence by offering evidence of something: a passport validates identity and a signed contract can have important practical consequences just as acts of speech can (Smith, 2014). Documents evince contexts beyond their immediate presence and, conversely, contexts frame the properties and powers of documents. These assumptions about documents ground and situate our discussion of context, relevance, and labor.

In what follows, we sketch the variety of ways that context has been approached in information science and, in Sections 2.1 and 2.2, describe how the word *context* provides a useful abstraction for organizing discussions of context by clarifying the great variety of connections and distinctions that can formulate the boundaries of contexts. In Section 3, we demonstrate the close relationship between context and relevance when considered through an abstract framework afforded by the words. Section 4 demonstrates through a thought experiment concerning information retrieval the integrated relationship between labor, relevance, and context within the framework we are proposing. We conclude by considering shared contexts, questions about what may lie beyond context, and the limits of our framework.

2 | CONTEXT

Agarwal's overview of how context has been approached in information science, *Exploring context in information behavior: Seeker, situation, surroundings, and shared identities* (2018), usefully enumerates many different ways in which information scientists have approached context. These include context as an environment or a container; as a setting; as a role; as a situation; as an actor's mind; as an information horizon, field, and pathway; as a constraint; as a life-world or information world; as common ground and ordinariness; as discourse; as information ground; as assigned meanings during interaction; as proximity and relevance; as time and place, embodiment and portability; and as legacy and determinant (pp. 8–18). This list does not exhaust the variety.

Lee (2011) suggests specific approaches to context tend to emphasize one or more of the following:

i. the set of symbolic expressions or representations that surround a [target entity, such as a word] and help one to express, make sense of, translate, or otherwise act upon or within it.
ii. objective or socially constructed characteristics and conditions of the situation in which a [target entity] is, appears or occurs.
iii. aspects of the mental or physical state, disposition, intentions, identity or recent experiences of a living actor that bear upon how she interprets, understands, acts within, or what she notices of, the situation at hand (p. 97).

Also worth noting is the distinction in user–computer interaction between the representation of a resource's context to assist retrieval and user behavior as a context (Dourish, 2004).

Even this longer summarizing list does not exhaust the variety. We could easily include cultural, interpersonal, intrapersonal, historical, geographical, institutional and many more contexts and approaches to context. Instead, we suggest how the word *context* can be used to organize the discourse about context.

2.1 | *Context* as an organizational tool

Examining this ever-expandable list of what has or could count as context in information science, it is clear that the word *context* as much as any shared empirical perspective organizes it. Thus, rather than attempting to define context as empirical phenomena, reify context as a psychological or social construction, or otherwise adopt a previously asserted conception of context, we investigate how the word *context* can help to provide a framework for considering the many kinds of context information science considers.

The origin of the English word *context* provides a useful way for considering context in information science

that does not require assumptions about the nature of locations, situations, or psychological states. *Context* is derived from the Latin *contextus*, meaning "connection," and *contexĕre*, which means "to weave together, connect" (Oxford English Dictionary, 2021, s.v. "context"). This focus on connections and woven relationships facilitates a framework for discussions of context that require fewer assumptions about what may constitute a situation or be "standing around" (circumstance). A weave can be complex, but it is ordered and organized by its fibers. Considering how these connections are enumerated enables a view of how contexts are articulated and can operate at different levels of hierarchy to shape the boundaries that define what has counted as information and what has counted as context. The theoretical implication of this formulation of context is that, without being reliant upon them, it can incorporate psychological states such as those described by Harter (1992) and by Sperber and Wilson (1986), for example, as well as the other kinds of context listed by Agarwal, Lee, and Dourish.

Use of the word *context* is an obvious choice for formulating a framework for approaching context, but other words can be similarly useful. For example, 脈絡 is a shared term for the idea of context in East Asia. This two-syllable compound is pronounced *maengnak* in Korean, *myakuraku* in Japanese, and *mailuo* in Chinese. Individually, the two sinographs suggest "(inter) connected (絡)" "veins (脈)." 脈絡 is associated with the idea of a circulatory system. Other words might also provide a useful way to think about the concept of context, but the English term *context* is a practical word with which to begin.

2.2 | Contexts and boundaries

Boundaries drawn between documents and their contexts are arbitrary. "Arbitrary" is used here to suggest subjective, considered judgment, rather than random states of affairs or, as in the case of linguistics, notions of intersubjectivity. As our long list of different kinds of context suggests, contexts can be expanded, contracted, moved, and rearticulated. Boundaries are descriptive because they indicate which items are associated with which other items. Like weaving, drawing boundaries is work. Like work, acts of judgment are related to purpose, which, like judgment, is prone to change. If drawing a boundary aids an enduring purpose, the boundary is likely to be maintained. For example, living organisms with perceptual systems that draw boundaries in the universe that aid their survival pass on their genes. Future generations then draw similar boundaries. In brief, boundaries between a document and a context are drawn arbitrarily as acts of judgment that serve a purpose.

Donna Haraway emphasizes the dynamic nature of these assessments, which she describes as "boundary projects":

> Boundaries are drawn by mapping practices; 'objects' do not pre-exist as such. Objects are boundary projects. But boundaries shift from within; boundaries are very tricky. What boundaries provisionally contain remains generative, productive of meanings and bodies. Siting (sighting) boundaries is a risky practice. (Haraway, 1991, p. 201).

Haraway means that objects can be formulated and reformulated according to how their boundaries connect them to, or distinguish them from, other objects. This dynamic helps to explain why documents can shape and be shaped by context. Since each interpretative judgment is both unique and dynamic, different individuals' boundary projects relating to the same document will necessarily differ but may well be related in potentially interesting ways.

Certain kinds of shared documents can be understood as "boundary objects," a term popularized by Susan Leigh Star (Bowker et al., 2015). Boundary objects are objects whose boundaries, that is, their connections and distinctions, have been stabilized so that they can be used as objects in and across communities to pursue a variety of ends. I may view a fence around my property differently from the way a neighbor views the same fence. But we share the fence as a physical boundary, which helps sustain the idea of a neighbor and of a neighborhood. Maps are never complete descriptions, but their incompleteness enables a variety of judgments ranging from navigational (which direction to go) to political and economic (which area to zone as residential).

That Haraway refers to objects as boundary projects suggests that they need never be finished and by implication any descriptive work to document an object, including documentary objects, is interminable. Objects as boundary projects can be connected with and distinguished from what comes before or after them, or above or below them. The objects of boundary projects blur into and emerge from what we have called context because the focus of attention is continuously changing according to how connections and distinctions are drawn and woven into other connections and distinctions.

Boundary objects are those objects whose connections and distinctions have been stabilized and delimited by use in a community. Boundary objects can be seen as integral to the formation of domains of knowledge. Birger Hjørland (2019, section 2.4) suggests, for example, that "a domain is a body of knowledge, defined socially and theoretically as the knowledge of a group of people sharing ontological and epistemological commitments." Domains

can be understood to contribute to contexts within the framework we are proposing but would not contain them. Domains, as the etymology of the word implies, have to do with control and the dominion shared ontological and epistemological commitments provide (often unequally) to individuals and communities. Contexts can, but do not necessarily, suggest specific ontological and epistemological commitments. Contexts, as we are considering them, are the connections (and distinctions) that boundaries make possible, the fences that individuate specific relationships (one family's property versus another's) and also broader communities (property owners versus non-owners). As Haraway and Star both remind us, drawing a boundary is an act of judgment with significant consequences. But contexts are not commitments to these or other ontological or epistemological formations. Contexts can enforce certain commitments but also require their abandonment.

3 | RELEVANCE, DOCUMENTS, AND AFFORDANCES

According to Tefko Saracevic, Samuel C. Bradford was the first writer to use the word *relevant* in an information science sense when, in the 1930s and 1940s, he wrote about articles relevant to a subject (Saracevic, 1975, p. 324). Surveying the literature on relevance between the 1930s and the 1990s, including Saracevic's, Harter (1992, p. 602) writes that "objective relevance is usually measured as topicality—how well the topic of the information retrieved matches the topic of the request. A document is objectively relevant to a request if it deals with the topic of the request." Topicality is critiqued by Harter, who proposes that "psychological relevance" (based on the ideas of Sperber and Wilson (1986) who studied verbal utterances) is more useful for information retrieval. Harter notes that others had critiqued the idea of topical relevance and proposed alternatives to capture the idea of what he calls "subjective relevance." These include Foskett (1972, pertinence), Kemp (1974, pertinence), Wilson (1973, situational relevance), Cooper (1973, perceived utility), Boyce (1982, informativeness), Buckland (1983, beneficiality), and Schamber et al. (1990, dynamic, user-oriented relevance). As with our list of kinds context in information science, Harter's list of kinds of relevance that have been studied in information science could easily be extended. Mizzaro (1997) and Saracevic (2007) enumerate additional approaches. In short, conceptions of relevance took new shapes with each "turn" in information science as these have been described by Hartel (2019): the "cognitive turn" of the 1980s; the "affective," "neo-documentalist," and "socio-cognitive" turns of 1990s, the "everyday life" and "social constructionist" turns at the turn of the new millennia, and the "embodied turn" in the mid-2000s.

Two things of note are revealed by this brief survey of relevance in information science. Like context, relevance has been of primary concern to researchers concerned with information retrieval. Second, these approaches all attempt to make claims on realities, whether those realities be subjective experiences that are called "relevant" or what are assumed to be real relationships between the topic of an information request and what is retrieved. In short, like context, relevance has been reified as an empirical phenomenon. We suggest the word *relevance* can be used as a productive tool for considering relevance. As with our choice of the word *context*, relevance is an arbitrary choice but one made based on our judgment that it can facilitate productive acts of judgment in information science and lessen work.

3.1 | Relevance and context

The noun *relevance* in general usage means "connection with the subject or point at issue; relation *to* the matter in hand" (Oxford English Dictionary, 2021, s.v. "Relevance"). Considering the semantic meaning of *relevance* reveals its close relationship with *context*. Both concern connections and, by extension, distinctions. Relevance, however, does not include all the connections to an issue or subject, but only those that are sufficient for some purpose or that make achieving some end easier. An older, now rarer, meaning is associated with Scottish law, where it suggests "legal sufficiency or adequacy." The etymology of the word's adjectival form suggests "easing, alleviating" and "to lighten."

This relationship between the word *context* and *relevance* facilitates a conceptual framework in which considering what is relevant in a context can be understood as a process of selective judgments. Not every possible connection that could be drawn to formulate a document and its context will also be connected to particular matters in hand. Many connections will be irrelevant to specific purposes. But others will be judged to alleviate work or to be adequate for some purpose. In other words, some strands of the weave that formulates a document and its context will have been judged to be apparently or potentially useful, which temporarily formulates a relationship between something considered potentially relevant and some reason or purpose for which it is judged to be relevant.

The relationship between *context* and *relevance* helps solidify a conceptual framework in which documents are characterized by a weave of connections and distinctions called context, which can be distinguished from connections and distinctions called relevance by what they

afford. The term affordance, as we describe below, denotes a capacity. It refers to what an object or idea enables while an end is pursued. Connections judged to afford some capability during the pursuit of some end can, in the framework we are proposing, be considered relevant.

The coining of the term "affordance" is often attributed to the psychologist James J. Gibson (1979). The concept, however, is older and was used by the Baltic German biologist Jakob von Uexküll (1864–1944). A biologist concerned with animal perception and behavior, Uexküll restricted his concept of the functional characteristics of objects to those that an individual animal could perceive. Fond of musical imagery, Uexküll used the term "functional tone" (in German *funktionale Ton*) for affordances that the living subject might find harmonious or discordant in relation to some purpose (Brier, 2008; Buchanan, 2008; Uexküll, 2010). This led to the influential concept of a subject's cognitive horizon (see Gadamer, 1985).

Because relevance can be conceptualized as a relation between an affordance and a purpose selected from all the connections that could be drawn to articulate a document and its context, a relevance judgment is articulated by context in at least three ways. The relevance *of* an affordance is relative to the relevance of other affordances available in the subject's context, and the relevance *to* a purpose depends on the unstable status of a purpose. Further, making relevance judgments requires more or less labor because, depending on the context, enumeration, comparison, and selection of potentially relevant affordances will take more or less work.

4 | LABOR

Labor is an exertion of energy in the mind, with the hand, or delegated to a machine. Some investment of labor is necessary to find affordances relevant to a given purpose. "Find" includes both discovery (identification) and access. The task of making a relevance prediction can be regarded as a mapping between two different kinds of goodness: quality and value. Quality concerns the capability of meeting some standard. How capably does an affordance serve an end? Value is a measure of benefit. What are the expected benefits of what is afforded? (Orr, 1973; Buckland, 1991, pp. 203–204).

4.1 | Two kinds of power

Patrick Wilson's *Two kinds of power: An essay on bibliographical control* (1968) imagines what he calls the bibliographical universe comprising all extant documents. He also imagines gaining control over it through description. What is needed, according to Wilson, is a means of gaining control over elements of the bibliographical universe so that they might better serve the ends people pursue with documents. He makes the case that bibliographical control is afforded by bibliographical description. He suggests two interdependent means of gaining control over what he would call a mass of mostly worthless and irrelevant things. He calls these interdependent means descriptive control and exploitative control. Exploitative control is Wilson's (1968) term "for the ability to make the best use of a body of writings" (p. 25). Descriptive control is his term for "an ability to line up a population of writings in any arbitrary order, to make the population march to one's command" (p. 25). A person with perfect descriptive control, Wilson suggests, can summon up every document that fits any desired description. A person with perfect exploitative control, "has merely to say what he wants writings *for*, and is then provided with what will suit his purpose best, whatever it is" (p. 25).

Exploitative power is relevant to some actual purpose. Descriptive power is, at most, relative to some imagined purpose that might never arise. Description establishes a ready-made context for making judgments about what is afforded among documents when a need arises. Exploitative power depends on a prior suitable investment in descriptive power, but future actual purposes can be predicted only imperfectly. The descriptions made may not be suitable and descriptive power that is never needed is a waste of descriptive labor.

Wilson's bibliographical powers suggests the important role played by various kinds of descriptive labor needed to articulate what can be considered context and what in any given context might be exploited as relevant affordances for a variety of pursuits.

4.2 | Precision and generality

A collection constitutes a context for the items within it. The labor of identifying and accessing particular documents in a collection will be affected by how documents are arranged within it. The challenge is to select only relevant documents. Various situations and purposes can be imagined and consequently, as we describe above, various elements of what is afforded by a document might be considered relevant. For simplicity here, we adopt the usual unrealistic assumption that relevance is binary: a document is either relevant or not relevant.

The selection labor required is very sensitive to the relationship between two very similar measures. In retrieval evaluation terminology, "precision" is the proportion of relevant documents in a retrieved set. High

precision reduces the wasteful inclusion of non-relevant documents. "Generality" is the proportion of relevant documents in the entire collection, which is usually extremely small. The retrieval task can be understood to be the deriving of a retrieved set with high precision from a collection with, typically, very low generality.

In a collection in which all documents were relevant to a given query, generality would be 1.0 (i.e., 100%). Successful retrieval would be trivially easy, and retrieval labor would be minimal because selecting any document at random would work as well as any other search strategy because, by definition, each retrievable document is relevant.

In a collection in which, say, only one in 10 documents is relevant, generality would be 0.1 (i.e., 10%). If the collection were searched at random one would expect only about one in 10 inspected items to be relevant. This would mean that precision would also be much reduced (to 0.1) and labor per relevant item retrieved would be much increased to an average of 10 documents inspected for each relevant item found.

For a more realistic scenario, imagine searching in a collection in which there is only one relevant document per ten thousand (generality 0.0001). Search at random would yield very low precision (0.0001) and an intolerable increase in selection labor: one relevant item per ten thousand inspected. Here we can use a relaxed sense of "random" to denote any ordering that does not aid selection. Imagine a telephone directory not in alphabetical order of name or the books in a library collection arranged by a subject classification when one wishes to search by author, publisher, or any characteristic other than subject.

Search labor increases as the proportion of relevant items in a collection decreases. Mathematically, the expected number of items inspected per relevant item found is the reciprocal of the generality. Search labor increases asymptotically with decreasing generality. Search at random or serial search in an unhelpful ordering would, of course, ordinarily be unrealistic because too laborious.

The efficient ideal is a collection in which relevant items are somehow pre-sorted to create an identifiable subset comprising all and only items relevant to one's current query. This theoretical ideal is not fully feasible because queries vary and are imperfectly foreseeable. Nor are queries a perfect proxy for purpose because the query has to conform to the searchable descriptions. Nevertheless, the challenge remains to improve selection performance from random retrieval toward perfect, efficient retrieval. The standard answer is to enrich the context by adding metadata. Adding descriptive metadata, for example, topic descriptors, enables the generation of indexes, which in theory could enable close to perfect retrieval. The author catalog for a library collection that is arranged by subject is a familiar example. As another example, consider finding words with a given meaning in a conventional dictionary in which words are arranged in alphabetical order. In both cases the labor needed to find desired entries would be prohibitively high, but the creation of an author index and the addition of topical indexing (as in Roget's thesaurus) respectively, would be transformative. In the first case, existing author statements have been indexed; in the second new topic descriptions have been added and used for the primary arrangement. The items sought have not been changed, but their context, that is, how they are connected and distinguished, has. The added metadata affords the necessary ingredients for the creation of contextualizing indexes, which constitute, in virtual form, a very large number of relevant subsets. If the indexing were perfect, the index headings would lead directly to sets of relevant documents with perfect precision. This simple thought experiment illustrates how selection labor depends on the context and how it can be dramatically transformed by a change of context.

4.3 | Two kinds of labor

Labor can be categorized as following rules (procedural) or as requiring judgments (arbitrary). We may imagine this to be a spectrum, but here, again for simplicity, we follow Julian Warner's "labor-theoretic" approach in which he distinguishes just two categories: semantic (arbitrary) labor and syntactic (procedural, algorithmic) labor (Warner, 2010, 2021).

Semantic labor concerns the work of making judgments about affordances without established, predetermined, or explicit criteria. It is the work of drawing connections that articulate a document and its context and also the work of judging which asserted connections afford some benefit and how much they afford to some purpose. This labor will vary depending on how a context has been enumerated initially but within a context is arbitrary. None of the relationships that are created and judged depend on an articulated rule. This framework for organizing context, relevance, and labor helps to explain why relevance and context have resisted scientific examination. Science requires agreed upon measures and boundary objects, whereas semantic labor is arbitrary. It is work done on evolving boundary projects.

Semantic labor can be contrasted with procedural, rule-based operations which, once defined, can be delegated to clerical or mechanical performance. Warner calls this procedure-based activity "syntactic labor." Syntactic

labor is premised upon established rules of description, of articulating what defines the connections between a document to its contexts and which connections are likely to afford some benefit to imagined purposes. As Wilson suggests (1968), a reciprocal relationship of some kind may be assumed but is not guaranteed between descriptive work and the utility of that descriptive labor while pursuing a bibliographical end. Investments in description, such as indexes, for example, reduce what Warner calls selection labor. They do so as much by limiting what can be selected among all the possible features of a collection as by formulating what counts as the collection. As Wilson points out, not all descriptions may be equally felicitous.

In the examples imagined above in the discussion of precision and generality, selection labor is reduced by 90% and by 99.9%, respectively if we exclude the one-time investment of descriptive labor required to generate the index. Perfect retrieval cannot be fully achieved in practice since, after Haraway, objects can be thought of as boundary projects and, hence, their description is interminable. Nevertheless, the reduction in labor through the use of contextualizing description can be very large. This is because description, like relevance, is selective. Indeed, only those connections deemed likely to be relevant to some real or imagined end are drawn. Through its selectivity, description generates a context in which most connections are likely to be relevant to certain kinds of ends, thereby reducing the labor of selecting among all the possible features that might afford a benefit.

Subjective judgments, including relevance judgments as they pertain to the use of descriptions but also the selective work of description, are asserted but are not ordinarily defined or explained. This has two consequences. First, semantic labor, like relevance judgments, resist scientific analysis. Second, surrogates for semantic labor may be very useful as proxies but inevitably remain unreliable. This is the case with the types of "objective relevance" already noted as convenient substitutes for a subject's subjective relevance judgments (e.g., the consensus of relevance judges). It is also the case for topicality as a useful predictor of what may be relevant to the inquirer even though topic and purpose are not the same. The same can be said of lexical similarity (the use of the same words or word-stems), a popular surrogate for topicality because it is so easy and inexpensive to generate using algorithms.

In Patrick Wilson's terms, descriptive power is the ability to orchestrate contexts through description such that the semantic labor needed to construct them is optimized to facilitate the reduction of labor needed when pursuing bibliographical ends. In Warner's terms, this is done by replacing semantic labor with syntactic labor. In conventional retrieval evaluation the emphasis has been on performance in terms of precision and recall of retrieved sets rather than the dramatic impact on cost and performance of contextual aspects. Warner's labor-theoretic approach draws attention to the dramatic cost implications of the type of labor deployed. Rather than assuming contexts as a priori states of mind or situations, the real opportunity and challenge is to ask how we might wish to orchestrate contexts such that labor toward certain ends are reduced. Warner's distinction between semantic labor and syntactic labor is useful. The former requires a great deal of energy since it is contingent upon fashioning a context that facilitates relevance judgments. The latter requires an initial investment in description but can be mechanized because it is based upon a previously defined context and set of rules for making relevance judgments.

5 | CONTEXTS AND PERCEPTIONS

So far attention has focused on the relationship between a document's affordances and its context, but how far does context extend? What is beyond it? And how might unique, different contexts be related or shared?

5.1 | Efficiency of perception

What, if anything, is beyond an individual's perceived context? A useful answer to this question is provided by a striking passage from Uexküll that describes the ability of a female tick to survive. The tick has limited sensory capabilities. She is blind, deaf, and has no sense of taste, but she can sense heat and smoothness and can smell butyric acid, which is associated with the sweat of mammals. The tick will climb a stalk or branch. When she smells butyric acid, she will leap off and land, if fortunate, on a passing mammal where she can find a smooth area of skin and suck blood (Uexküll, 2010, 44–45).

Uexküll coined the German term *Umwelt* (literally, surrounding world) to denote the tick's *perceived* environment or "perception world": the world as the tick can understand it. *Umwelt*, for Uexküll, meant all and only the affordances accessible to an organism through its own perceptual capabilities. Each creature has its own unique *Umwelt*, Uexküll argues, and each creature's *Umwelt* is a closed world. (*Umwelt* has since acquired a broader meaning in contemporary German as a general term for the environment.)

There is, of course, a lot more in the surroundings that a blind and deaf tick does not perceive. For this Uexküll used *Umgebung* (surroundings). A biologist

observing the tick would sense some of the surroundings that the tick cannot, but the biologist's range of perception (the biologist's own *Umwelt*) is also limited. The biologist also has unperceived surroundings (*Umgebung*) composed, for example, of items too small to see or outside the visible spectrum. Stated differently, every individual's cognitive context has a horizon limited by the individual's perceptual abilities.

The limitations of these perceptual contexts are crucial. Everything a subject perceives and every effect that it produces belongs to its perception world and forms a closed unit. This closed world has profound implications. It also enables a way to imagine what is not context: anything beyond a creature's *Umwelt* is imperceptible to a creature's perceptual systems and cannot be discerned. Uexküll writes that "the whole rich world surrounding the tick is constricted and transformed into an impoverished structure. However, the poverty of this environment is needful for the certainty of action, and certainty is more important than riches" (Uexküll, 2010, p. 51). The incompleteness of the tick's context productively alleviates its labor by efficiently presenting only that which is likely to be important for its survival.

This foray into the world of the tick, although perhaps unorthodox, helps to clarify our framework for considering context in information science as it relates to relevance and to labor. Uexküll suggests how contexts are formulated through what Wilson might call descriptive power so that certain kinds of relevance judgments about what has been afforded are easier because what Warner calls selection labor has been reduced. Uexküll's ideas about perception, interpretation, and "perception world" were of lasting interest to philosophers interested in hermeneutics, phenomenology, and "being" (Brier, 2008; Buchanan, 2008). In particular, Martin Heidegger's central concept of "being in the world" (*Dasein*) closely resembles Uexküll's "perception world."

5.2 | Different perceptions

Uexküll (2010, pp. 126–32) relates how an oak tree's affordances are differently perceived by different creatures. His description helpfully clarifies how contexts can be shared while remaining distinct within the framework we are proposing. According to Uexküll, an oak tree's roots afford a fox an opportunity to burrow and create a secure den. The oak's branches afford "handy springboards" for squirrels and places to land for birds. Bark that affords softness to the ichneumon wasp burrowing in to feast upon the larvae of bark beetles affords hardness in the environments of many other creatures. The enumerative sensory apparatuses of each species produce different contexts that articulate relevant means for their survival ends. In other words, they afford different means of impoverishing the extreme complexity of the world for their individual, species-specific purposes.

5.3 | Convergent perceptions

Similarities are to be expected in the separately perceived contexts of similar subjects of the same species or human communities because members will perceive in similar ways. Foxes perceive as foxes do, even if individual foxes see or hear things differently, for example. Ants perceive as their perceptual apparatuses permit, scientists as scientists do, and humanists as their perceptual orientations permit. In phenomenology, a "horizon" is, in general terms, a context in which any particular perception is situated. Inasmuch as shared understanding is taken to involve a "fusion of horizons," it always involves the formation of a new context that enables integration of what is otherwise unfamiliar, strange or anomalous. So intersubjective understanding involves mediation and dialogue (Malpas, 2018). Gadamer discusses this convergence of perception using the visual horizon figuratively for a cognitive context. "To acquire a horizon means to look beyond what is close at hand—not to look away from it, but to see it better within a larger whole and a truer proportion," he noted. Changing perceptions mean changing horizons: "Horizons change for a person who is moving" (Gadamer, 1985, pp. 271).

6 | CONCLUSIONS

This view of horizons and our foray into the environments of animals provide a useful frame for concluding our consideration of how context, relevance, and labor can be related in information science, in part by providing an unorthodox context. They suggest a way to think about how context has been approached in information science and a means to suggesting the infinity of ways that context might be considered. Our suggestion is that fields within information science have been working within their own environments. What is called context by those in information retrieval is distinct to the environments of information retrieval and beyond the horizon of archivists, for example. The framework we propose promises a view of the oak tree that supports all of information science's communities and the variety of contexts within which they work. We extend Uexküll's metaphor to emphasize that the framework we propose is not a description of information science as any particular reality but a means for organizing the diverse plurality of ways that information scientists generate and consider

contexts, facilitate relevance judgments, and labor to alleviate work. Our proposal is that concepts suggested by the words *context* and *relevance* afford a descriptive framework, and thus some descriptive control, of the diversity of contexts and kinds of relevance formulated in information science. While this initial attempt at descriptive control will be incomplete necessarily, it can be exploited to reduce the amount of work we must do to find and make use of different approaches to context and relevance in information science. As we suggest in our introduction, the connections and distinctions that *context* enables holds out the possibility that what counts as context in information retrieval can inform those working on preservation and information theory and vice-versa.

To summarize, we propose a framework in which contexts can be considered the sum of connections (and distinctions) drawn into a weave that articulates the boundaries of documents and the horizons beyond them. Relevance judgments can be understood as judgments made by a subject about the affordances of certain connections within a context while pursuing a specific end. Labor can be described as a function of the effort needed to make relevance judgments as they are afforded by a context. To facilitate the creation of this framework, rather than assuming contexts to be a priori material circumstances, social situations, or cognitive dispositions, we formulated context and relevance abstractly by making use of what is afforded by *context*, *relevance*, and their etymologies. Our framework does not exclude empirical work of information scientists on the complex realities of context, of course. Rather, it aims to provide a theoretical framework for organizing its diversity by providing a new horizon for considering context, relevance, and labor.

ORCID

Wayne de Fremery https://orcid.org/0000-0002-2716-8335

REFERENCES

Agarwal, N. K. (2018). *Exploring contexts in information behavior*. Morgan & Claypool.

Blair, A., Duguid, P., Goeing, A. S., & Grafton, A. (Eds.). (2021). *Information: A historical companion*. Princeton University Press.

Bowker, G. C., Timmermans, S., Clarke, A. E., & Balka, E. (Eds.). (2015). *Boundary objects and beyond: Working with Leigh star*. MIT Press.

Boyce, B. (1982). Beyond topicality: A two stage view of relevance and the retrieval process. *Information Processing and Management*, *18*(3), 105–109.

Brier, S. (2008). *Cybersemiotics: Why information is not enough!*. University of Toronto Press.

Buchanan, B. (2008). *Onto-ethologies: The animal environments of Uexküll, Heidegger, Merleau-Ponty, and Deleuze*. SUNY Press.

Buckland, M. K. (1983). Relatedness, relevance and responsiveness in retrieval systems. *Information Processing and Management*, *19*(4), 237–241.

Buckland, M. K. (1991). *Information and information systems*. Praeger.

Buckland, M. K. (2014). Documentality beyond documents. *The Monist*, *97*(2), 179–186. http://escholarship.org/uc/item/55v7p74x

Buckland, M. K. (2017). Document theory. In B. Hjørland (Ed.), 2021 *ISKO encyclopedia of knowledge organization*. International Society for Knowledge Organization. http://www.isko.org/cyclo/document

Cedars Project Team. (2001). T*he cedars project report April 1998 – March 2001*. Consortium of University Research Libraries http://web.archive.org/web/20040623020330/http://www.leeds.ac.uk/cedars/pubconf/papers/projectReports/CedarsProjectReportToMar01.pdf.

Cooper, W. S. (1973). On selecting a measure of retrieval effectiveness. *Journal of the American Society for Information Science*, *24*(2), 87–100.

de Fremery, W. & Buckland, M. K. (2022). Copy theory. Journal of the Association for Information Science and Technology, 73 (3), 407–418. http://dx.doi.org/10.1002/asi.24558

Dourish, P. (2004). What do we talk about when we talk about context? *Personal and Ubiquitous Computing*, *8*(1), 19–30.

Fetzer, A. (2017). Context. In Y. Huang (Ed.), *The Oxford handbook of pragmatics* (pp. 259–276). Oxford University Press.

Foskett, D. J. (1972). A note on the concept of 'relevance'. *Information Storage and Retrieval*, *8*(2), 77–78.

Gadamer, H. G. (1985). *Truth and method*. Crossroad Publishing Company.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.

Haraway, D. J. (1991). *Simians, cyborgs, and women: The reinvention of nature*. Routledge.

Hartel, J. (2019). Turn, turn, turn. *Proceedings of CoLIS, the Tenth International Conference on Conceptions of Library and Information Science, Ljubljana, Slovenia. Information Research*, *24*(4), paper colis1901. http://InformationR.net/ir/24-4/colis/colis1901.html

Harter, S. (1992). Psychological relevance and information science. *Journal of the American Society for Information Science*, *43*(9), 602–615.

Hjørland, B. (2019). Domain analysis. Version 1.4. In B. Hjørland (Ed.), *ISKO encyclopedia of knowledge organization*. International Society for Knowledge Organization. https://www.isko.org/cyclo/domain_analysis

Hjørland, B., (Ed.), (2021). *International Society for Knowledge Organization encyclopedia of knowledge organization*. https://www.isko.org/cyclo/

Kelly, T. (2016). Evidence. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy (winter 2016 edition)*. Stanford University https://plato.stanford.edu/archives/win2016/entries/evidence/

Kemp, D. A. (1974). Relevance, pertinence and information system development. *Information Storage and Retrieval*, *10*(2), 37–47.

Lee, C. A. (2011). A framework for contextual information in digital collections. *Journal of Documentation*, *67*(1), 96–143.

Malpas, J. (2018). Hans-Georg Gadamer. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy (fall 2018 edition)*. Stanford

University https://plato.stanford.edu/archives/fall2018/entries/gadamer/

McDonald, J. D., & Levine-Clark, M. (Eds.). (2018). *Encyclopedia of library and information sciences* (4th ed.). CRC Press.

Mizzaro, S. (1997). Relevance: The whole history. *Journal of the American Society for Information Science*, *48*(10), 810–832.

Orr, R. M. (1973). Measuring the goodness of library services. *Journal of Documentation*, *29*(3), 315–332.

Oxford English Dictionary. 2021. Online Oxford University Press. https://www.oed.com/

Pagès, R. (2021) Documentary transformations and cultural context (M. K. Buckland, trans). *Proceedings from the Document Academy*, 8(1). https://ideaexchange.uakron.edu/docam/.

Saracevic, T. (1975). Relevance: A review and a framework. *Journal of the American Society for Information Science*, *26*(6), 321–343.

Saracevic, T. (2007). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: Nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology*, *58* (13), 1915-1933; Part III: Behavior and effects of relevance., *58*(13), 2126–2144.

Saracevic, T. (2017). Information science. In M. J. Bates & M. N. Maack (Eds.), *Encyclopedia of library and information sciences* (4th ed., pp. 2216–2231). Taylor and Francis.

Schamber, L., Eisenberg, M. B., & Nilan, M. S. (1990). A reexamination of relevance: Toward a dynamic, situational definition. *Information Processing and Management*, *26*, 755-776.??

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3) 379-423, 623–656.

Smith, B. (2014). Document acts. In A. Konzelmann-Ziv & H. B. Schmid (Eds.), *Institutions, emotions, and group agents. Contributions to social ontology* (pp. 19–31). Springer. http://ontology.buffalo.edu/smith/articles/document-acts.pdf

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Blackwell.

Stock, W. G., & Stock, M. (Eds.). (2013). *Handbook of information science*. De Gruter Saur.

Uexküll, J. v. (2010). *A foray into the world of animals and humans*. University of Minnesota Press.

Warner, J. (2010). *Human information retrieval*. MIT Press.

Warner, J. (2021). *Copyright, data, and creativity in the digital age*. Routledge.

White, H. D. (2018). Relevance in theory. In J. D. McDonald & M. Levine-Clark (Eds.), *Encyclopedia of library and information sciences* (4th ed., pp. 3926–3939). CRC Press.

Wilson, P. (1968). *Two kinds of power: An essay on bibliographical control*. University of California Press.

Wilson, P. (1973). Situational relevance. *Information Storage and Retrieval*, *9*(8), 457–471.

How to cite this article: de Fremery, W., & Buckland, M. K. (2022). Context, relevance, and labor. *Journal of the Association for Information Science and Technology*, *73*(9), 1268–1278. https://doi.org/10.1002/asi.24631