A Rich Context Model for Knowledge-Works

Arijit Laha
C-KDIS, Infosys Technologies Ltd.
Hyderabad

Abstract

Lack of context in information is a serious problem for knowledge-workers. Effective utilization of computational aids for supporting knowledge-workers require a rich understanding of the nature of context of information and related knowledge-works. It also needs specifications about how such understanding can be leveraged in computer-based systems. In this paper we propose a holistic model of context of knowledge-works and information created in course of their performances. We also demonstrate with an example how such a model can be used as basis for developing a formal, machine-deployable specification of activity context.

Introduction

The computing research community is currently in the process of investigating various aspects of ”Context-aware” (Dey 2001) computing. The notion of ”context” is ancient, overarching and highly multi-faceted one (Dourish 2004). Depending on the class of computing applications being considered, researchers usually use the term in much narrower senses. For example, researchers of mobile and ubiquitous/pervasive computing largely focus on spatial and temporal contexts. On the other hand, for Information Science researchers (Mizzaro 1997), context is anything other than document metadata and explicit system parameters (e.g., query) which contributes in establishing relevance of archived information. See (Kaenampornpan and O’Neill 2004) for a tabular comparison of several definitions of context.

Here, we examine the notion of context from the perspective of designing a class of computer-based systems/applications, aimed at aiding/assisting human agents to perform complex ”knowledge-intensive activities” or ”knowledge-work” of professional nature. These activities are exemplified by various forms of design, research, planning, decision-making, etc. in domains of professional activities (e.g., governance, finance, health-care, information technology, etc.). An study on some of the desired characteristics of such systems and nature of the problems faced in building such systems can be found in (Markus, Majchrzak, and Gasser 2002). Nevertheless, state-of-the-art of computer science and technology holds out great promises to overcome many of those problems to a significant extent.

A deep and rich understanding of the nature of “context” and its role in knowledge-works and means to represent them in a formal manner, is necessary for harnessing the power of modern computational techniques in this area. To this end, in the rest of the paper first we examine the nature of interaction between human cognition and information with respect to a knowledge-work. Then we propose a model of *rich context*, interlinking *activity*, *information* and *work-environment*. In the next section we provide an example of how the proposed context model can be formalized and used in for building an advanced computing environment for providing holistic support to knowledge workers.

The subject matter of this paper is in use as the basis of an ongoing research project. Many of the ideas and concepts described here has been implemented in a prototype system (Laha 2010). The results of preliminary studies conducted with the prototype are fairly encouraging. Nonetheless, there is a need for further research into different facets of the theoretical foundation described here as well as on possible ways it can be exploited.

Knowledge-work and Information

The nature and relationships among the ideas of “knowledge”, “information” and consequently, “knowledge-work” is a subject of enormous and still ongoing debate involving a number of disciplines. Thus, the positions developed here may be treated as “working definitions/hypotheses” for serving the current purpose rather than claims of universal nature.

According to Polanyi (Polanyi 1967), “knowledge” is tacit and thus personal to the “knower”. Information is produced when a knower “articulates” symbolically (speaking, writing, gesture ...) some explicit/explitizable part of her knowledge. Then, *symbolic information* becomes available to others for interpreting and understanding, with the help of their individual *mental models*, which, in turn, consists of the ideas, assumptions, beliefs, facts, even misconceptions, which together shape one’s worldview (Johnson-Laird 1983). This consumption of information leads to gain/update of their individual possessions of knowledge. We can derive two interesting implications from the above.

- Knowledge is not sharable in whole; and
- From the same information, different persons gain non-identical knowledge due to their difference in mental models.

A knowledge-work is undertaken in order to *solve a problem*. Bohme and Stehr (Böhme and Stehr 1986) viewed a knowledge-work as the site of *(re)production of knowledge and information*. Production of knowledge corresponds to a knowledge-worker accessing relevant information, interpreting and understanding them. This leads her to conceive a solution to the problem-at-hand or gaining some useful insight into the nature of the problem and/or its solution(s). Information is produced/created when the worker articulates the newly gained knowledge in symbolic form. Thus, from the perspective of *computer-based assistive system*, which deals essentially with information, a knowledge worker can be supported while she is engaged in "information usage", i.e.,

consuming information, i.e., trying to seek, retrieve and interpret/understand relevant information; and

(re)producing information, i.e., articulating newly gained knowledge and recording/capturing the resulting new information.

Clearly, in order to achieve consistency in results and their quality while performing same/similar types of knowledge-works, the workers need to possess significant degree of similarity in relevant aspects of their mental models (Alavi and Leidner 2001). This is achieved by means of suitable education, training and experience of the workers. However, apart from these long-term factors, during performance of a knowledge-work, a worker needs to incorporate in her mental model the *context of current work* or simply "work-context". While consuming an information she (re)constructs the context of the information, from various available cues, and compare it with her own work-context. Again, while producing information, she attempts to embed various cues, explicit (e.g., document metadata, references) and implicit (logical/semantic structure of discourse), which might help future consumers to reconstruct the current context.

Context model for knowledge-work

So, what constitute the "context" of information with respect to information usage in a knowledge-work? The question assumes great importance with on-going information explosion and consequent information overload faced by knowledge-workers. To combat these problems researchers have built systems like ASAP (Glasner et al. 2006), Codex (Pike and Gahegan 2007), AWARE (framework) (Bardram and Hansen 2010) etc. in specific types of knowledge-intensive activities in particular domains of works.

Clearly, there is a growing demand for such systems. Thus, at this point, it is important to investigate issues related to formulation of *a general framework/architecture* for designing and building advanced context-aware computing environments for supporting knowledge workers. However, a vital prerequisite to develop such a framework is a deeper and holistic understanding of the relevant elements of "context" and their interrelations for knowledge-work space. Only then we can successfully bring together the combined power of modern computing technologies to the aid of knowledge-workers in tackling the problem of capturing and reconstructing context during information usage.

On the outset, we observe the following:

- Information is tangible result/outcome of performance of knowledge-intensive activities (for the sake of simplicity, at this point we are not concerned with automatically collected data, e.g., sensor data, or even data generated due to routine transactional/operational activity) performed by human actors;
- A knowledge-worker's work-environment supports her in performing a knowledge-work;
- Elements (including cognitive and social ones) of work-environment are utilized and/or modified by the knowledge-worker in course of her performance;

Now, let us define the *context of a knowledge-work*, or *work-context*, for brevity, as consisted of the relevant elements of the knowledge-worker's work-environment. Also, let us call the activity where a particular information originated, the *source activity* of the information. Then, we can define the potentially persistent and sharable (part of) *context of an information* as consisting of *information about*

- the work-context of its source activity; and
- how the elements of work-context were utilized in the course of its performance.

Thus, a "context model for knowledge-works" needs to identify relevant elements of a knowledge-work environment, their interrelations and how they are utilized in course of its performance. Then only we can proceed to formulate means for capturing, associating and exploiting the information about them for helping in information usage processes. In order to arrive identify relevant elements of knowledge-work environment, first we look up to more general notion of "human activity".

Human activity

A human activity is an immensely complex, multi-faceted interaction of human cognition and elements of its environment. A class of theories, e.g., distributed cognition, actor-network theory, activity theory, collectively referred as "post-cognitivist theories" (Kaptelinin and Nardi 2006), rooted in psychology and other social sciences, attempts to provide a holistic perspective of this interaction. The viewpoint described below is *inspired* by the "activity theory (AT)", especially as interpreted in (Kuutti 1996).

To suit our limited purpose, i.e., creating a *context-model for knowledge-works* which can be used *to organize information in computer-based systems*, we conceive that a human activity takes place when a "human actor" attempts to achieve some "objective". Achievement of an objective results in creation and/or transformation of some entities, the "outcome", which may be tangible, often physical objects such as a house or a machine and/or abstract entities such

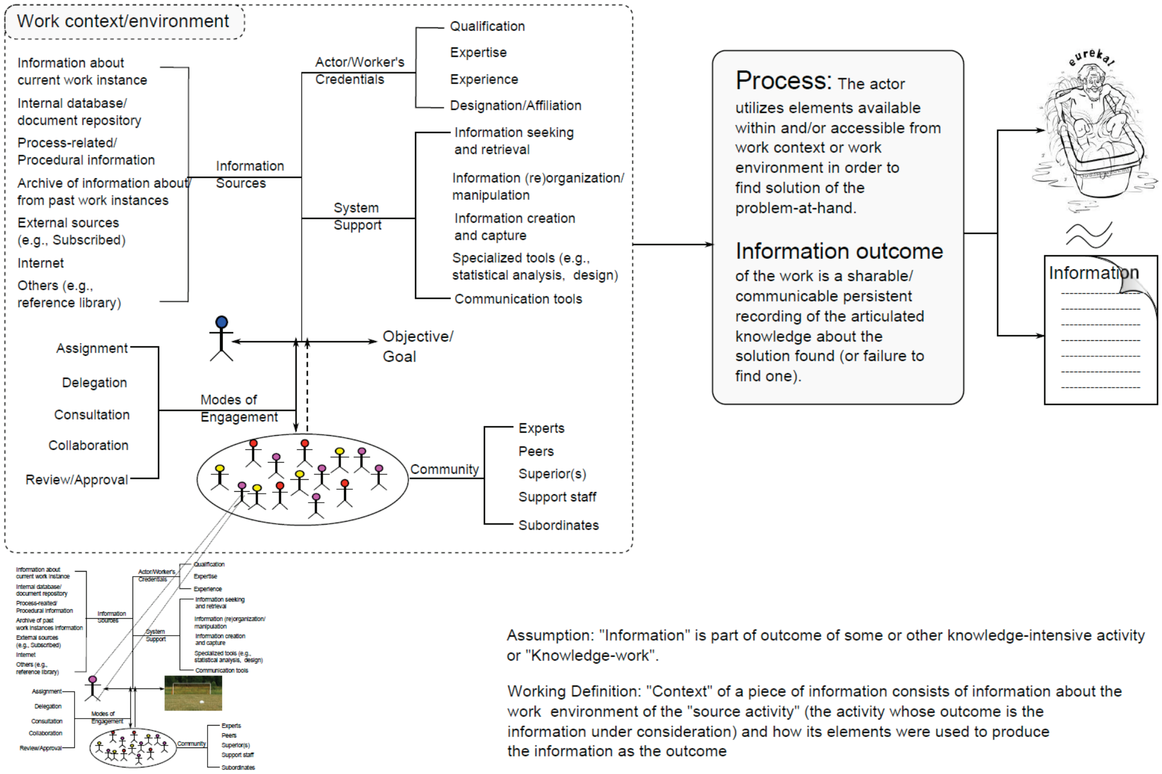

Figure 1: Elements of a knowledge-work

as knowledge, experience, etc. To be able to achieve the objective, the actor uses various kind of resources, including knowledge, experience, physical tools, help from others and so on. All these available resources are utilized by the actor in suitable manner while she is enacting a “process” that takes her gradually closer to achievement of the objective.

We also borrow from AT the notion of levels of activity. AT (Kuutti 1996) recognizes three levels of activity, *activity*, *action* and *operation*. The exact boundary of these levels and whether they are enough for succinctly analyzing human activity, is a subject of some debate (Kaptelinin and Nardi 2006). However, for our purpose we shall use these three levels as structural units of activity for which we shall define the distinguishing properties.

Knowledge-intensive activity

If we specialize the above viewpoint with respect to knowledge-works, we can identify various elements involved in a knowledge-work as shown in Figure 1. Figure 1 depicts the top level categories of the elements are identified, each of which may be treated as a root of a hierarchy of subcategories. In the following we describe the elements, with the help of examples from a *patient-care* activity in which a physician and other staff in a hospital treats an ailing patient.

Actor A knowledge-work is performed by a human actor (indicated by the stick-figure in Fig. 1), e.g. a physician in a patient-care activity.

Objective The objective or goal of a knowledge-work is to solve a problem that requires intellectual/cognitive contribution. It can be creation of some tangible artifact, e.g., a design document, and/or something abstract, as in the case of patient-care it is to bring back the patient to a state of good health.

Resources A knowledge-worker needs a diverse set of resources. We divide them into three main categories,

1. Actor’s credentials, i.e., qualification, expertise, experience and designation/affiliation. While these are mostly intangible, information about them help establishing authenticity of the information produced;
2. Information sources available to the actor, both internal and external;
3. Supporting systems (mostly computational for our purpose) for dealing with information.

Community A knowledge-worker’s community is comprised of people to which she can turn for help in achieving the object, whether mandatorily (e.g. the physician asking a pathologist to run a test on the patient) or optionally (the physician consulting a colleague about the condition of the patient). Here we are assuming that the role and contribution of a community member is function of her credential and mode of engagement (Fig. 1).

Note that, like the actor, every community member also has her own work-environment. Thus, every engagement is an interaction of two (or more) work-contexts. This has several implications, which we shall leave out of the scope of the current paper, in possible uses of the context-model in collaborative/cooperative setup.

Process The process of a knowledge-work represents the actor's progress towards achievement of the objective by suitable utilization of elements of work-context. The knowledge of the process may be possessed by the actor intellectually and/or it may be available in various codified forms as part of the system/tools. In the next section we provide an example of the later case.

Outcome Outcome of a knowledge-work can be divided in two related parts. Due to her intellectual involvement, the worker gains knowledge, not only about the solution of the problem, but also adequacy or otherwise of the work-context as well as strengths and weaknesses of the process used. More important from an information-processing system perspective, outcome includes the information produced by the worker through articulation of the gained knowledge.

The context-model

Thus, for a knowledge-work, actor, objective, tools/resources, community and possible modes of engagements of its members together constitute the *work-environment*, and in turn, its *work-context*. Also, since an information is *an outcome of its source activity*, its context include information about the work-context and information about the process which led to its production as outcome.

Above definition and its depiction in Figure 1 provides us with an abstract context-model. This can be instantiated for a particular knowledge-work by mapping suitable domain specific elements onto the model. These elements may include taxonomies, domain ontologies, etc., enabling a wide gamut of useful computational techniques. For example, one can use a context-model along with a text analysis technique in order to categorize segments of a text document and compute its relevance to the current work-context.

However, in order to capture and exploit the information and its context in sufficient detail by computational means, we need to develop a method of formalizing the context-model. This can take several form, depending on the techniques (first order logic, semantic network, petri-net, process calculus, etc.) used and the granularity level of activity required to be supported by the application. In the following section we describe a graph-theoretic technique.

Formalizing the context-model

The Knowledge-work Support System (KwSS) (Laha 2010) is a design framework of a class of assistive system for knowledge-workers. However, here the term "KwSS" also refers to a system (e.g., the prototype described in (Laha 2010)) built using the framework. In the following we, shall use "a KwSS" to refer to a system while "the KwSS" will indicate the framework.

The KwSS incorporates a crucial idea, based on human of cognitive limitations, that actual performance of a knowledge-work, and hence, consumption of information takes place at a granularity level of *cognitively manageable complexity*. Consequently, the production of information and its context is best achieved at the same level, temporally close to the gain of knowledge. Thus, given a complex "task", the largest granule of activity a KwSS is designed to support, an activity model is incorporated into the system that seamlessly maintains the context-model up to the desired level of granularity. This allows a KwSS, at proper granularity levels, to:

- Assist a worker to easily access elements of work context, through various modes of computing facilities, e.g.,
  - Process-related/procedural information through guidance;
  - Easy access to relevant information;
  - Systematic interactions with relevant community members (technically, interactions among different work-contexts);
- Assist a worker to create new information and capture them easily along with their detailed context.

Activity Modeling

An "Activity Model (AM)" for a knowledge-intensive activity expresses (1) the environment of the activity; (2) interdependency among the constituent sub-activities; and (3) information as outcome of the activity and its context in form of argumentative supports, references and annotations.

Activity representation An activity (knowledge-work) a_i is formally represented as a triple

$$a_i =< E_i, P_i, O_i >$$

in the model, where,

- E_i is a set consisted of the (hierarchies of) elements of the work-context as depicted in Figure 1;
- O_i is the information outcome of the activity; and
- P_i represents the process (to be)enacted by a knowledge worker to produce O_i.

For an activity a_i, sufficiently complex, such that it needs to be decomposed into a number of smaller (granule of activity) sub-activities, P_i is a decomposition graph (V_i, E_i), representing the set of sub-activities V_i and their interrelations. Such an activity is called a "composite activity".

In contrast, an activity a_i, that can be accomplished by performing a set of actions, is called a "Simple Activity". Here, an **action** is an activity where the human actor needs to *consciously choose and apply a sequence of operations* in order to achieve the desired outcome. (e.g., selecting a search tool use, building a query string, running the tool with the query as its input/argument), while an **operation** represents an activity that is supported (automated) by the available system(s) to such an extent that the actor need not be aware of the details about how the underlying process takes place (e.g., firing a query to a search tools and getting the output). In other words, for a simple activity a_i, the process P_i is not a decomposition graph, but a *sequence of actions*.

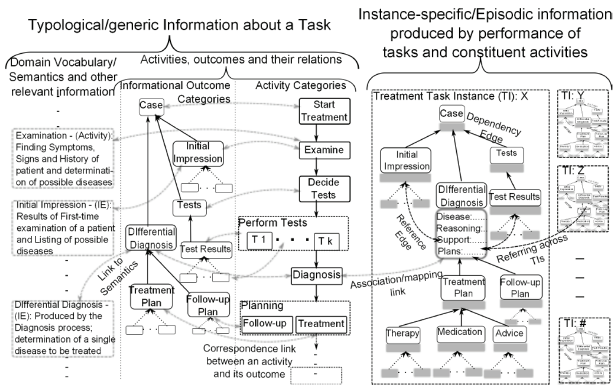

Figure 2: Associations among elements of TM and EM

Typological and Episodic activity models With respect to KwSS design, we distinguish between the *typological* and *episodic* structures and hence models of an activity. While both have same structural property, they represent two different classes of information.

The "Typological Model (TM)" of an activity-type represents the typological or categorical information about various components of the activity and elements of its environment. The TM of an activity can be prepared off-line and deployed in a KwSS as a "reference" or "nominal model" in order to provide various types of sophisticated support to a knowledge worker while performing an episode/instance of the activity-type. Such supports include:

- Guidance though a reference process structure;
- Quick access to relevant part of available knowledge-base (ontologies, thesauri, reference documents); and
- Enabling context-aware deployment of advanced computational techniques (e.g., information filtering, recommendation, semantic comparison, inferencing, etc.).

An "Episodic Model (EM)", in contrast, corresponds to a particular episode (or instance) of performance of an activity. It represents the information about the *activity as performed* and the *information as produced* along with their contextual associations in course of performance. In other words, The EM is the episodic information along with its context. Elements of EM, where their typologies are available in the TM, are associated with them and other semantics/referents through it. Each EM is archived in order to capture and retain episode-specific richly contextualized information, as they are developed during the performance of the episode. The correspondence patterns between TM and EMs in a KwSS is depicted in Figure 2.

Formal codification of "process"

Codification of a *process* involve formal definition of its structure and semantics of its elements. Here we demonstrate the codification of the process in "complex activities" in form of a graph-theoretic formalism.

Structural properties An activity $a_i =< E_i, P_i, O_i >$ can be represented graphically as depicted in Figure 3.

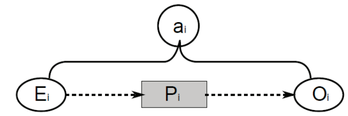

Figure 3: Graphical representation of an activity

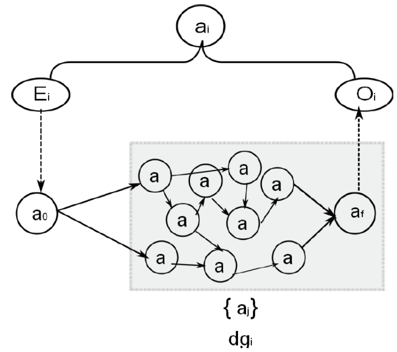

Figure 4: The decomposition graph

For a composite activity a_i, the process P_i has a decomposition structure represented by a decomposition graph $dg_i = (V_i, L_i)$ with the following properties:

1. $V_i = \{a_j\} \cup a_0$ where, $a_j =< E_j, P_j, O_j >$ is a subactivity of a_i and $a_0 =< E_i, \emptyset, E_i >$ is a special (dummy activity);
2. There is one and only one activity $a_f \in V_i$, called the final activity, such that $O_f = O_i$;
3. An edge $l_{jk} = a_j\vec{a}_k \in L_i$ is a directed edge from a_j to a_k and represents *dependency* of a_k on a_j;
4. The set of activities $\{a_j|(a_j \in V_i) \wedge \exists(l_{jk} \in L_i)\}$ is called the *dependency set* of activity a_k, denoted as $DSet(a_k)$;
5. The activity a_0 does not have any *in-edge* and the activity a_f does not have any *out-edge*;
6. Each node $a_k \in V_i$ belongs to *at least one directed path* from a_0 to a_f, i.e., $\forall a_k \in V_i,\ \exists(a_0 \rightsquigarrow a_k \rightsquigarrow a_f)$; and
7. There exists *at least one topological ordering* of V_i.

While the properties 1 to 4 are descriptive one, the rest (5-7) are constraints which ensure that

DAG the dg_i is a Directed Acyclic Graph (DAG); and

Dependency each node, except a_f belongs to the dependency set of at least one other node, i.e.,

$$\forall a_j \in (V_i - a_f),\ \exists a_k \in V_i\ such\ that\ a_j \in DSet(a_k).$$

The structure of dg_i is graphically depicted in Figure 4.

Model semantics The structural formalism described above, provides us with the means to define a formal semantic model of activity and information that follows the proposed context-model (Fig. 1). The formal sematic model of an activity in from the KwSS perspective can be described as follows:

- Given the representation of an activity $a_i =< E_i, P_i, O_i >$, the component
  - E_i represents a systematic access to all the resources available to a worker for performing the activity;
    - E_i evolves with the progress of the activity through accumulation of information produced by its different constituent parts (sub-activities, actions, operations);
  - O_i represents the information outcome, both in form and content;
  - P_i is the process involved in producing the outcome, which, in case of a simple activity is a sequence of actions and otherwise a decomposition into system of smaller activities;
- In a system of activities, some activities require availability of one or more particular pieces of information those are not available in the initial environment. Production of these information must also constitute the part of the larger(super) activity. This dependency is represented in the dg_i through the construct of dependency set. Thus, formally:
  - $\{O_j | a_j \in DSet(a_k)\} \subset E_k$,
    where, $a_j \in DSet(a_k) \Rightarrow \exists(l_{jk} \in L_i)$;
- The activity a_0 represents the transfer of resources available at super-activity to the sub-activities;
- The activity a_f represents the final sub-activity whose completion indicates *completion* of the super-activity.

Above specification of activities is used in KwSS framework to design assistive systems for a specific, target knowledge-work type. While a knowledge-worker is using the system to perform an instance of the target type, the system aids her in context-aware manner by organizing and coordinating access to relevant resources as well as allow her to easily record context, maintained to a large extent by the system, along with information.

For example, while a physician is trying to diagnose a patient and queries the case archive for information about a disease, the system automatically prioritize the cases where the disease was diagnosed rather just mentioned in different context (e.g., as part of medical history). To achieve such capabilities, the system uses the model to leverage computational techniques such as semantic analysis/comparison, text analysis, recommendation, etc. more effectively than is possible otherwise. Also once she makes and records the diagnosis, the system automatically associate with it relevant data points and references used, annotations/observations made by her as basis of the diagnosis.

Conclusion

In this paper we have proposed a rich context-model for knowledge-works and information created in course of them. There are many possible ways of utilizing the model for supporting aspects of knowledge-work. Here we demonstrated that through an example where the context-model is used for developing a formalism. The formalism can be used for designing advanced supports for knowledge-workers by coordinating resources in a context-aware manner as well as marshalling the various advanced computational techniques in more effective ways.

References

Alavi, M., and Leidner, D. E. 2001. Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly* 25(1):107–136.

Bardram, J. E., and Hansen, T. R. 2010. Context-based workplace awareness. *Comput. Supported Coop. Work* 19:105–138.

Böhme, G., and Stehr, N. 1986. *The Knowledge Society*. Springer. chapter The Growing Impact of Scientific Knowledge on Social Relations, 7–30.

Dey, A. 2001. Understanding and using context. *Personal and Ubiquitous Computing* 5(1):4–7.

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

Glasner, J.; Rusch, M.; Liss, P.; III, G. P.; Cabot, E. L.; Darling, A.; Anderson, B. D.; Infield-Harm, P.; Gilson, M.; and Perna, N. T. 2006. Asap: a resource for annotating, curating, comparing, and disseminating genomic data. *Nucleic Acids Research* 34(Database issue):D41–D45.

Johnson-Laird, P. N. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Boston: Cambridge University Press.

Kaenampornpan, M., and O'Neill, E. 2004. Modelling context: An activity theory approach. In *in Ambient Intelligence: Second European Symposium, EUSAI 2004*, 367–374. Springer.

Kaptelinin, V., and Nardi, B. A. 2006. *Acting with Technology*. Cambridge, Mass.: MIT Press.

Kuutti, K. 1996. *Context and Consciousness: Activity Theory and Human Computer Interaction*. MIT Press. chapter Activity Theory as a Potential Framework for Human-Computer Interaction Research, 17–44.

Laha, A. 2010. On the issues of building information warehouses. In *ACM Compute 2010, January 22-23, Bangalore, India*.

Markus, M. L.; Majchrzak, A.; and Gasser, L. 2002. A design theory for systems that support emergent knowledge processes. *MIS Quarterly* 26(3):179–212.

Mizzaro, S. 1997. Relevance: The whole history. *Journal of the American Society for Information Science* 48(9):810 – 832.

Pike, W., and Gahegan, M. 2007. Beyond ontologies: Toward situated representations of scientific knowledge. *International Journal of Human-Computer Studies* 65(7):674–688.

Polanyi, M. 1967. *The Tacit Dimension*. New York: Anchor Books.