WITTGENSTEIN, LANGUAGE AND INFORMATION

# Information Science and Knowledge Management

## Volume 10

# WITTGENSTEIN, LANGUAGE AND INFORMATION

'Back to the Rough Ground!'

*by*

DAVID BLAIR

*University of Michigan, Ann Arbor, Michigan, USA*

Springer

This book is dedicated to my sons, Alain and Christopher.

# Table of Contents

# Acknowledgements

My first book, Language and Representation in Information Retrieval, was so well received that to write a book that was an extension of its theories was a difficult and time-consuming process. I spent 10 years writing this book which extends the analysis of my first book. During that time I greatly appreciated individuals who read my manuscripts and gave me their impression of what I had written. The most dedicated reader of my manuscripts was John Klinkert from Seattle, Washington, who read each of several versions of this manuscript, and sent me written collections of his comments. I'm currently a professor at the University of Michigan. The Michigan faculty member who read early versions of my book was Kevin Kerber. His comments were very supportive and helpful. In Chicago, former University of Chicago professor, Donald Swanson read several versions of my manuscript, and was supportive enough to invite me to his house to discuss what I had written. I also found some of his own publications relevant to my analysis, so they are cited and discussed in the final manuscript.

One of the most supportive acts concerning my early manuscripts of this book was performed by Hubert Dreyfus, professor of philosophy at the University of California at Berkeley. He read an early version of my manuscript and then quoted several paragraphs of it in his new book, On the Internet (Routledge, 2001). This is especially complementary because Hubert Dreyfus is one of the premier philosophers who uses philosophy to analyze practical problems. His 1972 book What Computers Can't Do was an analysis of the potential of Artificial Intelligence, and has proven to be an extremely accurate predictor of the limitations of it. (It's in its 3rd edition now, and has been renamed What Computers (Still) Can't Do, MIT Press) Overall, Dreyfus is the most inspiring modern philosopher for my work in applying philosophy to the practical problems of information systems as I did in this book.

Other individuals who read and responded to early manuscripts of this book were Richard Boland, professor at Case Western University, M.E. Maron, professor emeritus of the University of California, Berkeley, Brian O'Connor, professor at North Texas University, Ron Day, professor at Wayne State University, and Eric Edderer, a professional librarian.

## Major Cited Works by Ludwig Wittgenstein and The Abbreviations Used in This Discussion:

| Works by Wittgenstein: | Abbreviations: |
|---|---|
| Blue and Brown Books | BB |
| Culture and Value | CV |
| Last Writings on the Philosophy of Psychology, v. 1 | LWPP I |
| Last Writings on the Philosophy of Psychology, v. 2 | LWPP II |
| Lectures and Conversations on Aesthetics, Psychology and Religious Belief | LC |
| Notebooks 1914–1916 | NB |
| On Certainty | OC |
| Philosophical Grammar | PG |
| Philosophical Investigations | PI |
| Philosophical Remarks | PR |
| Remarks on the Foundations of Mathematics | RFM |
| Remarks on the Philosophy of Psychology, v. 1 | RPP I |
| Remarks on the Philosophy of Psychology, v. 2 | RPP II |
| Tractatus Logico-Philosophicus | TLP |
| Zettel | Z |

# Part I: Introduction

> "The more narrowly we examine language, the sharper becomes the conflict between it and our requirement. (For the crystalline purity of logic was, of course, not a result of investigation; it was a requirement.) The conflict becomes intolerable; the requirement is now in danger of becoming empty.—We have got onto slippery ice where there is no friction and so in a certain sense the conditions are ideal, but also, just because of that, we are unable to walk. We want to walk; so we need friction. Back to the rough ground!"[1]
>
> —Ludwig Wittgenstein

This manuscript consists of four related parts: a brief overview of Wittgenstein's philosophy of language and its relevance to information systems; a detailed explanation of Wittgenstein's late philosophy of language and mind; an extended discussion of the relevance of his philosophy to understanding some of the problems inherent in information systems, especially those systems which rely on retrieval based on some representation of the intellectual content of that information. And, fourthly, a series of detailed footnotes which cite the sources of the numerous quotations and provide some discussion of the related issues that the text inspires. The first three of these parts can each be read by itself with some profit, although they *are* related and do form a conceptual whole. Still, the reader who wants an overview of many of the arguments advanced herein, can get them comparatively quickly from Part I, while the reader who wants to see, in some detail, the exegesis of Wittgenstein's late philosophy of language and mind would do well to read Part II with some care. Of course, the central message of the manuscript is presented in Part III, where the implications of Wittgenstein's later philosophy for information systems, especially information retrieval systems, are worked out in some detail, providing a deeper discussion of the issues described in Part I. The only part which cannot be read by itself is, of course, the footnotes, what I would call Part IV. Footnotes have an ambivalent status in writing. Manuals of style insist that if the material is important enough to be included in a manuscript it should be placed in the text and not in the footnotes; if it's not important enough to include in the text, it's not important enough, with few exceptions, to be in the footnotes either. Some individuals are even more pointed in their dislike of footnotes. John Barrymore once said "A footnote in a book is like a knock on the door downstairs while you are on your honeymoon." Certainly footnotes *can* interrupt the flow of the manuscript, and long footnotes can take the reader far enough away from the discussion of the main text that it may be difficult to return to it. Yet, some texts demand the links and extended discussion that footnotes provide, and this present text is one of them. In the first place, any detailed exegesis of Wittgenstein's, admittedly difficult, philosophy, requires citations to identify the wide variety of sources of the frequent quotations and the related or alternative versions of cited passages. Wittgenstein's philosophy, in particular, has generated an enormous amount of critical analysis since his death half a century ago, and to ignore the major points of this analysis would be unwise, so references have been made to discussions

---

[1] **Philosophical Investigations**, §107, 3rd ed. Blackwell Publishers, Oxford, UK, 2001. Translated by G.E.M. Anscombe. 1st ed. published 1953. [Hereafter referred to as **PI]**

or critiques of his work, especially those made by the premier Wittgenstein scholars, G.P. Baker and P.M.S. Hacker, and Wittgenstein's former student and professor of philosophy, Norman Malcolm, who spent his career as professor of philosophy at Cornell University.[2] Wittgenstein himself was not a systematic writer, developing many of the main themes of his work in a sporadic fashion throughout his writings. This unsystematic development of his work demands that the selections from his writings which are used here be carefully cited so that the reader can see where, in the 15 volumes of his currently published writings, the exact quotations come from. Because Wittgenstein's later philosophy is spread out across so many individual works, it would be unreasonable to expect the reader to have copies of all of Wittgenstein's writings for reference. Consequently, I have made every attempt to quote relevant passages as completely as possible, and, in some cases, to give examples of the variations of the same statement in different parts of his writings. To this end, the reader will not be required to have a library of all of Wittgenstein's works cited herein in order to follow the thread of this discussion.

Another reason for the large number of detailed footnotes is the nature of the subject of this discussion—language. The subject of language is as broad and deep a topic as there is. Language permeates almost every aspect of our lives, and the nature of meaning resists concise or comprehensive explanation. Language, it can be argued, is part of the very definition of what it means to be human, and any discussion of the nature of language inevitably brings up myriad links to the cognitive, social and cultural dimensions of mankind. One cannot write succinctly about the nature of language and meaning and still hope to capture its depth and complexity. As Wittgenstein himself put it, "...words have meaning only in the **stream of life**."[3] This is not the first text on language to deal with its complexity in this way, Noam Chomsky's first two, enormously influential, books on language, **Aspects of the Theory of Syntax** (1969) and **Cartesian Linguistics** (1966) had similar styles.[4] Both of these works brought out the complexity and depth of the subject of language by providing detailed footnotes that greatly increased the length of the texts themselves. Any serious discussion of language use and its related issues is, fundamentally, an intellectual adventure. The footnotes in this text are meant to describe the many directions and dimensions of this adventure. Certainly, if the pleasures of the text are enough to make the reader regret the distraction of a footnote, then the best advice is to keep reading and, in John Barrymore's metaphor, simply refuse to answer the knock at the door. As I pointed out above, the text in large part does stand by itself.

## Why Language?—Why Philosophy?—Why Wittgenstein?

Making our way through Wittgenstein's late philosophy of language and mind is a demanding journey, so it will be helpful to first survey Wittgenstein's intellectual landscape

---

[2] The most extensive bibliography of Wittgenstein criticism is **Wittgenstein: A Bibliographical Guide**, by G. Frongia and B. McGuinness [Basil Blackwell, Oxford, UK, 1990]. This work lists references to 1,942 articles and texts published between 1914 and 1987 which discuss Wittgenstein's work. Many of the references have brief annotations.

[3] **Remarks on the Philosophy of Psychology,** v. II, §687. The University of Chicago Press, Chicago, 1980. Edited by G.H. von Wright and H. Nyman, translated by C.G. Luckhardt and M.A. Aue. [Hereafter referred to as **RPP II]**

[4] **Aspects of the Theory of Syntax**. MIT Press, Cambridge, MA, 1969. **Cartesian Linguistics: A Chapter in the History of Rationalist Thought**. Harper Collins, NY, 1966.

to identify its major features and how they relate to the issues of information retrieval. This will give our later analysis a clearer foundation. Wittgenstein wrote frequently of the importance of the "Übersicht," or "Overview" of examples necessary to see how a word is used (see the section in Part II "Five Red Apples" for a more detailed discussion of the German word "Übersicht"). This brief survey of some of his ideas and how they relate to information systems provides a kind of "Übersicht" of his philosophy of language and will set the principal features of the intellectual landscape on which we will make our explorations in Parts II and III.

First of all, why are the issues of language and meaning important to the study of information systems? Information systems are, of course, tools that are used to search for information of various kinds: data, text, images, etc. Information searches themselves inevitably require the searcher to ask for or describe the information he or she wants and to match those descriptions with the descriptions of the information that is available: in short, when we ask for or describe information we must *mean* something by these statements. This places the requests for information as properly within the study of language and meaning.[5] Surely, requests for information, or descriptions of available information, can be clear or ambiguous, precise or imprecise, just as statements in natural language can. In short, understanding how requests for, and descriptions of, information work, and, more importantly, how they can go wrong, is an issue of language, meaning and understanding.

Why, then, is the focus of this discussion on philosophy? Certainly, the fields of linguistics and literature, especially literary criticism, have much to tell us about language and meaning. I would agree, and should I write another book I might be tempted to look to those fields for enlightenment about meaning and language. But I'm turning to philosophy of language first for the principal reason that its *main* concern is with how we *mean* what we say—how does language actually work? These are important issues in linguistics and literary study, but they are not the *central* concerns of these fields. Since the problem of meaning in language is the central concern of this discussion, its aims and focus most closely parallel the aims and focus of the philosophy of language. Another reason why the philosophy of language is particularly pertinent for the present discussion is that for philosophy in general, and Wittgenstein in particular, there is no sharp boundary between understanding language and cognition—how we understand language is closely coupled with how we understand things in general. Not only language, but understanding is important for information systems, too, since information systems are often used to help us understand things better. As we will see in Part III, information systems are part of what I, following Clark,[6] would call the "scaffolding" of our thought. I take the approach of philosophy of language to be the fundamental examination of the issues of meaning, so if there are any clear insights into our understanding of meaning, they will likely be found here first. This is why the philosophy of language is so important to our investigation. Nevertheless, we must be selective about what we use from the philosophy of language. Like any other intellectual discipline, the philosophy of language has its own specific

---

[5] A brief introduction to the relation between information retrieval and the philosophy of language can be found in "Information Retrieval and the Philosophy of Language," by D. Blair [**Annual Review of Information Science and Technology**, vol. 37, pp. 3–50, 2003.]

[6] A. Clark. **Being There: Putting Brain, Body, and World Together Again**. MIT Press, Cambridge, MA, 1997.

puzzles—what Thomas Kuhn called "exemplars"[7]—that preoccupy its practitioners. Some of these are helpful in our investigation, and some are not. Looking at the conduct of the philosophy of language in general, one can discern a number of courses its study has taken: early work, primarily that of Bertrand Russell and Gottlob Frege, was concerned with the relationship between language and the world. For them, language was primarily used to make factual assertions about the world. The central question of this era was "What are the truth conditions of a statement?" that is, "What is the relationship between meaning and truth?" While these are important questions for the philosophy of language, and are still prominent concerns of some present day work, the truth conditions of a statement do not tell us much that would be important to the study of language and information. In recognition of this, our discussion must be selective about what we use from the philosophy of language, and very thorough about extracting the maximum benefit from those aspects of philosophy that are relevant to our present study. "Ockham's Razor" is no less relevant to our present study than it was to the 14th century philosopher William of Ockham in his defense of Nominalism. Another obvious question must be, why is the philosophy of Wittgenstein particularly important for this study; that is, why not just survey the pertinent sections of the Philosophy of Language in general? There are many philosophers of language, and many philosophical theories which have contributed to our understanding of meaning in language. Why should we concentrate our efforts on Wittgenstein's, admittedly difficult, philosophy of language? Surely there are other, easier, routes to furthering our understanding of language and meaning. But Wittgenstein is unique among philosophers in the following respect: early in his career he was the consummate logician, the intellectual heir apparent to the pioneering logical work of Gottlob Frege and Bertrand Russell. Frege and Russell believed that ordinary language was not precise enough to represent the complexity and subtleties of meaning that were becoming increasingly important for analytic philosophy. Russell believed that the goal of analytic philosophy was to clarify what we say about the world. Analytic philosophy should take its inspiration from what Russell believed was the rigor of the scientific method. Since different branches of science often needed their own representational systems to express factual scientific relationships clearly, philosophy would need a similar rigorous representational system to make what it could assert perfectly clear, or so Russell and Frege thought. What we needed, they believed, was a logical language that could faithfully model these complexities and subtleties of expression, and could be used to clarify whether statements of fact were true or false—a language that could be used to bring out and make explicit the underlying logic of language. Early in his career, Wittgenstein was sympathetic with this view of language, believing, like Russell and Frege, that language could be made more precise through the use of formal logic. In his introduction to Wittgenstein's first published work, **Tractatus Logico-Philosophicus**, Russell describes Wittgenstein as being "concerned with the conditions which would have to be fulfilled by a logically perfect language."[8] Russell goes on to describe a logically perfect language as one which "has rules of syntax which prevent

---

[7] T. Kuhn. **The Structure of Scientific Revolutions,** 2nd ed. University of Chicago Press, Chicago, 1970. See the "Postscript" in which he discusses the "disciplinary matrix" of which exemplars are a part.

[8] Bertrand Russell: Introduction to the **Tractatus Logico-Philosophicus**, page ix. Routledge and Kegan Paul, London, 1961. English translated by D.F. Pears and B.F. McGuinness. First German edition published in 1921. [Hereafter referred to as **TLP]**

nonsense, and has single symbols which always have a definite and unique meaning."[9] But as Wittgenstein's thought matured, he began to have serious misgivings about the ability of logic to model or represent the complex and subtle statements of language. Not only was logic inadequate to this task, he thought, ordinary language itself was, if used properly, the best possible medium for linguistic expression, philosophical or otherwise. In short, Wittgenstein's thought evolved from a belief that problems of meaning in language could be clarified by logically analytical methods to a realization that many of the unclarities of language were a result of removing statements from the context, practices and circumstances in which they were commonly used—what Wittgenstein called our "Forms of Life." What determined the truth or meaning of a statement was not some underlying logic, but how the statement was used and what circumstances it was used in. Ambiguities in language are clarified, not by logical analysis, but by looking at how the words or phrases in question are used in our daily activities and practices. Wittgenstein's transition in his view of language is important for the study of information systems for the following reason: our current most widespread model of information systems is the computer model, in particular, the "data model" of information. This has been a very successful and robust model that has had a remarkably long history of implementation. Computers are, in a fundamental sense, logical machines, so we might say that the current most popular model for information systems is the *logical model*. This logical model, as we will show, has worked well for providing access to the precise, highly determinate content of our data bases—things like names, addresses, phone numbers, account balances, etc. But as more and more of our information is becoming managed by computerized systems we find that we must provide access to less determinate information, like the "intellectual content" of written text, images, and audio recordings—for example, searching for information that analyzes the economic prospects of Central European countries, or information that evaluates the impact of government regulation on small businesses. These kinds of access are not as well served by the logical data model of information, as one can easily see when trying to find some specific subject matter (intellectual content) on the World Wide Web using an Internet search engine.[10] Current information systems are in some way, the victims of the success of the more determinate data model of information. The logical/data

---

[9] *Op.cit.*, p. x.

[10] I first made this distinction between the more precise logical model of Data Retrieval and the less precise, more problematic model of Document Retrieval in "The Data-Document Distinction in Information Retrieval." [**Communications of the ACM**, vol. 27:4, pp. 369–374, April 1984.]

Some readers will no doubt rejoin that they have no trouble finding the intellectual content they are seeking on the WWW, and, in a certain sense, this is true—one can always find something relevant to one's search on the web. The key issue here is not finding "something relevant," but finding the best information for some purpose available on the web. Swanson described some time ago what he called the "Fallacy of Abundance":

> A scientist who nowadays imagines either that he is keeping up with his field or that he can later find in the library whatever may have escaped his notice when it was first written is a victim of what might be called the "fallacy of abundance." The fact that so much can be found on any subject creates an illusion that little remains hidden. Although library searches probably seem more often than not to be successful simply because a relatively satisfying amount of material is exhumed, such success may be illusory, since the requester cannot assess the quantity and value of relevant information which he fails to discover. [D.R. Swanson. "Searching Natural Language Text by Computer," **Science**, vol. 132, pp. 1960–1104, 21 October 1960. Quotation p. 1099.]

model of information has become the Procrustean Bed to which many information systems are forced to fit.[11] The effort to fit language and information to the logical model was justified because it was assumed that, as Russell and the early Wittgenstein believed, there is an underlying logic of language that governed its correct usage—an underlying logic which must be uncovered if we wanted to insure the clarity of expression. On this view, information systems used to provide access to "intellectual content" are just sloppy or imprecise versions of data retrieval systems. But it was one of Wittgenstein's clearest reassessments of his early philosophy when he said that "... the crystalline purity of logic was, of course, not a result of investigation; it was a requirement"—that is, the logic that Russell and Frege sought to uncover in their analysis of language, did *not* exist latently in language waiting to be uncovered. The logic of language was something that was a requirement for the analysis *to begin with*—it was something that was imposed on language. Just as Wittgenstein began to have misgivings about the applicability of the logical model, with its requirement for the strict determinacy of sense, to all aspects of language and meaning, some, this author included, are now having misgivings about how applicable the logical/data model of information is to the more complex and subtle problems of access to less determinate information such as the "intellectual content" of written text, images and audio recordings, a kind of access becoming increasingly widespread as more and more of our information starts out in machine readable form. As Douglas van Kirk put it:

> Corporations everywhere are beginning to recognize that information is almost always document-oriented. Because so many companies are in the information business, it stands to reason that the most productive companies will be those that manage documents effectively.

The reason why documents are so important to organizations is that they provide the context that makes information more meaningful. Wittgenstein, the consummate logician, came to see the limitations of logic when used to analyze language, and tried during the remainder of his career to indicate what was rigorous and right about ordinary language. Given that language *does* make sense, how does it do this without the armamentarium of logic? Wittgenstein's answer to this is both relevant to, and important for, access to intellectual content. Ordinary language is good enough for our purposes:

> It is wrong to say that in philosophy we consider an ideal language as opposed to our ordinary one. For this makes it appear as though we thought we could improve on ordinary language. But ordinary language is all right. Whenever we make up "ideal languages" it is not in order to replace our ordinary language by them; but

---

This was a remarkably prescient insight for 40 years ago, at the very beginning of the explosive growth in available information, and the electronic revolution in information storage and retrieval. Swanson could see this affect of abundance even in the research libraries of that time. Those systems, of course, are dwarfed by the size of today's World Wide Web, with its billions of separately searchable and retrievable pages.

[11] One needs only to look at the advertisements of the largest data base management system manufacturer, Oracle, to see this.

They advocate storing *every* kind of information, data, text, audio, images, video, on their systems, with no indication that each of these information types might require vastly different access methods.

> just to remove some trouble caused in someone's mind by thinking that he has got hold of the exact use of a common word.[12] [**BB** p. 28]

> On the one hand it is clear that every sentence in our language "is in order as it is." That is to say, we are not striving after an ideal, as if our ordinary vague sentences had not yet got a quite unexceptionable sense, and a perfect language awaited construction by us.—On the other hand it seems clear that where there is sense there must be perfect order.—So there must be perfect order even in the vaguest sentence. [**PI** § 98]

Wittgenstein is not saying that we are never misunderstood when we use ordinary language; of course we are. Wittgenstein is just clarifying how we should identify and resolve these misunderstandings. Instead of building a "logically perfect language" that would be more precise than our day-to-day language, or using logical methods to analyze linguistic mistakes, we must reorient our investigation: Instead of looking for an *underlying logic* of language, we need to look at how language is *actually used*, for it's not an underlying logic that clarifies what we mean, it's the context, activities and practices in which we use language that provide the fundamental clarification of meaning we are looking for. This is why Wittgenstein's work is so relevant to the study of information systems. Formal logic is useful for clarifying or solving a narrow range of problems and puzzles in language. Russell's, and the early Wittgenstein's, mistake was to think logic was applicable for solving a wide variety of linguistic problems that went beyond this narrow range. Like formal logic, the data model of information systems is an enormously successful model for the design of a narrow range of information retrieval tasks—those systems which provide access to highly determinate information (names, addresses, phone numbers, account balances, etc.). The data/logical model of information is a less successful model for providing access to the less determinate "intellectual content" of things like documents and images. The data/logical model cannot always capture the subtleties of language necessary for the retrieval of precise intellectual content on large information systems (again, searching for specific intellectual content on the World Wide Web is a good example). And, like language, there is no underlying logical model of information that we need to uncover— the "crystalline purity of logic" for information systems, like language, is "not a result of investigation; it [is] a requirement." That is, for the data/logical model to be applicable to all information systems, it is *required* that the information on the system be represented in extremely precise or determinate ways. But this process will have the effect, not of making better, "more precise" information systems, but, in the case of the search for "intellectual content," of making dysfunctional information systems—systems which are insensitive to the subtleties of language that are required for highly specific access to intellectual content, especially on large systems. As long as we believe that the precision of representation for data retrieval is possible *for all information systems*, we will run the risk of building such dysfunctional systems.

But recognizing the problems with the logical/data model of information is only part of the problem. What is needed to replace the inadequacies of the logical/data model?

---

[12]**The Blue and Brown Books,** p. 28. **The Blue and Brown Books** were published as a single work by Basil Blackwell, Oxford, UK, 1958. Cited edition is Harper Torchbook, New York, 1965. [Hereafter referred to as **BB]**

Wittgenstein's reassessment of the logical model of language, and his assessment of how language *really works*, is, it is the thesis of this discussion, a good guide for how the logical/data model of information systems must be changed or evolve if it is to provide satisfactory access to less determinate information such as intellectual content. Most philosophers of language have not been logicians, and those who had expertise in logic were not logicians of the first rank, as Wittgenstein was. This is why his reassessment of the usefulness of the logical model of language carries so much weight, and is particularly relevant to the present study. One of the central issues of language, for Wittgenstein, was what he called the "determinacy of sense"—the precision by which meaning can be defined. As we will see, it reappears as a fundamental issue in information retrieval too.

## Surveying Wittgenstein's Landscape

To begin our discussion, it will be useful to provide a brief overview of some of the major themes of Wittgenstein's philosophy of language and mind, and their relation to information retrieval systems. These major themes can be represented by carefully selected quotations from his works.

> 1. ... we don't start from certain words, but from certain occasions or activities.
> [**LC** p. 3][13]

Language does not exist by itself in a static system of definitions and syntax, but is intimately caught up in our activities and practices, what Wittgenstein called our "forms of life."[14] These forms of life comprise what Wittgenstein referred to as the "common behavior of mankind."[15] Language is not so much a collection of "meanings" but something that can be used to *do* things—it is an essential part of our everyday activities and practices. This makes meaning a largely collective notion: meaning *emerges* from the *use* of language in the conduct of day-to-day activities and practices. Emergent phenomena occur when a broad, higher level pattern emerges from the personal interactions of individual entities in the absence of any central controller. In language, "meaning" emerges from the interactions of native speakers using language in their day-to-day activities.[16] A dictionary definition

---

[13] **Lectures & Conversations on Aesthetics, Psychology and Religious Belief**, p. 3. University of California Press, Berkeley, 1972. Edited by C. Barrett. [Hereafter referred to as **LC**]

[14] "... the *speaking* of language is part of an activity, or of a form of life." [**PI** §23]

[15] "The common behavior of mankind is the system of reference by means of which we interpret an unknown language" [**PI** §206]. This is the reason that, as Wittgenstein comments:

> "If a lion could talk, we could not understand him." [**PI** p. 223]

Even if the lion were to use the same vocabulary as we do, because his day-to-day activities are so different from ours, he would use our words in strikingly different ways. He would also have no conception of the specific uses of words which we consider commonplace (e.g., a lion would have no idea what a "pet" is because pets have no role in its day-to-day activities, and what he would consider "food" would be strikingly different from our own conception). If our speaking lion were to say he was "stepping out for lunch" we could be assured that he was not going to a restaurant.

[16] Zipf asserted that the fundamental linguistic interaction which governs language use is the "competition" between speakers and hearers, where the speaker is trying to express himself as economically as possible,

does not *precede* the use of an expression, but *emerges* from the day-to-day use of the expression as a component in a particular activity or practice. As Wittgenstein put it, "...words have meaning only in the stream of life" [**RPPII** §687]. There are no private linguistic meanings or private languages—my sensations are, of course, personal and cannot be felt by others, but while my pain may be personal, I can only talk about it using the expressions commonly used to talk about pain.[17]

The idea of "emergence" is a relatively recent idea, but Wittgenstein seems to have anticipated some of the basic characteristics of this phenomenon. John Holland[18] discusses how the process of emergence can give the impression of something orderly and meaningful arising out of "chaos" or "disorder." Wittgenstein hinted at this possibility of order arising from disorder:

> No supposition seems to me more natural than that there is no process in the brain correlated with associating or with thinking; so that it would be impossible to read off thought-processes from brain-processes. I mean this: if I talk or write there is, I assume, a system of impulses going out from my brain and correlated with

and the hearer is, himself, trying to exert as little effort as possible to understand him. The "competition" occurs because the speaker tries to use as few words as possible, thus minimizing his effort, while the hearer is simultaneously trying to get the speaker to use more words in order to simplify his efforts at understanding the speaker. Zipf calls this the competition between speaker and hearer, or the competition between the forces of "unification" and "diversification." Optimal linguistic understanding occurs when the efforts of the speaker and hearer reach a balance where the sum of their respective efforts is a minimum. It is from these local interactions between speakers and hearers that meaning in language emerges. (for a more detailed presentation of Zipf's theory of language see the section "Implications of the 'Language as City' Metaphor" in Part II).

[17] Wittgenstein's argument against private languages or meanings rests on two assertions: first, private languages can only be discussed or described using the ordinary public language which is available to us all. But if the meaning of private language is expressible only through public language then it is not a private language at all, but merely a somewhat different form of ordinary, public language. As Wittgenstein put it:

> "What goes on within ... has meaning only in the stream of life" [**Last Writings on the Philosophy of Psychology**, vol. II, p. 30. Blackwell Publishers, Oxford, UK, 1992.]. [Hereafter referred to as **LWPP II**]

Secondly, language requires exemplary uses of words as guides to meaningful usage. We do the same kind of thing when we suggest that certain paintings are exemplars of a particular style. In art we can say that if you want to look at an exemplar, or, good example, of impressionist painting look at Monet's work. In language, similarly, we can say that if you want to know when to use the word "charisma" think of John F. Kennedy at a presidential press conference, or if you want to know how to use the word "compassion" think of Mother Teresa's work with the desperately poor—these are exemplary cases. But since the meanings of words in a private language would be entirely personal, you would not be able to establish such exemplars—that is, since, in a private language, a word can mean anything you want it to, there could be no examples which are better than others—in other words, there could be no role for exemplars to play. Even if you could establish exemplars in a private language, how would you know whether you were applying the exemplars correctly or not? The essential interplay between speakers and hearers, by which we gauge the correctness of our usage, and from which correct usage emerges, would be missing. There are no criteria for the establishment of exemplars in a private language, in fact there are no coherent criteria for the correct usage of any words of a private language. Since the speaker and the hearer are the same individual in a private language, there is no chance of there ever being a misunderstanding. Again, for Zipf, meaningful language arises from the "balance" achieved by the competition of Speakers and Hearers, yet in a private language there is no such interplay and no "balance" of usage, hence, there can be no private linguistic meaning in the ordinary sense.

[18] J. Holland. **Emergence: From Chaos to Order**. Helix Books, Addison-Wesley, Reading, MA, 1998.

my spoken or written thoughts. But why should the system continue further in the direction of the centre? Why should this order not proceed, so to speak, out of chaos?[19]

**Relevance to Information Systems:** The underlying order of information systems, in so far as they are linguistic systems, is not so much words and categories, but "occasions and activities." Yet often it is the relation to "occasions and activities" that is lost when information is organized for retrieval, especially when it is placed on a computerized retrieval system. Consider a simple example. Paper-based information has some obvious disadvantages regarding storage and copying when compared to the same information in electronic form. But paper-based information has one distinct advantage over electronic information: since a paper document does not need delicate electronic equipment to present it, it can be carried and used almost anywhere, from the office, to the home, to a bus, to a rainy construction site, etc. It is also easy to mark up, annotate or highlight paper, and parts of it can be clipped out or xeroxed and distributed. Further, small accidents such as dropping the paper or spilling coffee on it do not render it unreadable, though information on a laptop computer could not stand such abuse. Consequently, paper-based information can remain close to the activities that produce or use it, and these activities can provide an interpretive context for that information. But when that information is computerized, the very act of computerization may have the effect of removing the information from the activity context that provides much of its meaning and interpretation.[20]

> 2. The best example of an expression with a very specific meaning is a passage in a play.[21]

We learn our native language not so much by memorizing the definitions of words and phrases, but by watching it being used, trying to use it ourselves, and having new expressions or subtleties of meaning demonstrated to us by others. The best definition of a new word or expression is not a dictionary definition, but a scene in a play. For Wittgenstein, descriptions or demonstrations (e.g., plays) are better ways of conveying or clarifying meaning than explanations (dictionary definitions). In fact, "stage setting" can be an essential component of meaning even for a linguistic act as simple as giving something a name:

> When one says "He gave a name to his sensation" one forgets that a great deal of stage setting in the language is presupposed if the mere act on naming is to make sense. [**PI** §257]

A dictionary definition is a kind of shorthand explanation, what Wittgenstein was later to call a kind of "Language Game," and typically can only help us if we already understand the general role of the word in language. When we understand a new expression, what we have is not the ability to recall its definition, but the ability to *use* the expression in

---

[19] **Zettel**, §608. University of California Press, Berkeley, 1967. Edited by G.E.M. Anscombe and G.H. von Wright, translated by G.E.M. Anscombe. [Hereafter referred to as **Z**]

[20] The importance of the proximity of information systems to the activities and practices they serve was a major theme of Blair's **Language and Representation in Information Retrieval**. [Elsevier, 1990]

[21] **Last Writings on the Philosophy of Psychology, v. I**, §424. University of Chicago Press, Chicago. Edited by G.H. von Wright and Heikki Nyman, translated by C.G. Luckhardt and Maximilian A.E. Aue. [Hereafter referred to as **LWPP I**]

the appropriate circumstances and context. We may be able to use many words correctly without being able to define them—Modern English, the language since Shakespeare, existed until the18th century without dictionaries, and it was many more years before dictionaries were widely available. Prior to the wide use of dictionaries, all language learning had to be done through the personal interactions of native speakers in, or with reference to, the relevant daily activities or circumstances. Yet these daily interactions, unaided by dictionaries, produced the rich and nuanced language of Shakespeare's plays and a wealth of literature, essays, history and philosophy. If dictionaries were not widely available until fairly recently, then, it seems that before that we had no central or "essential" criteria for correct meaning or usage.[22] But there is no evidence that the lack of dictionaries caused any great confusion in communication: There are other criteria for correct usage, criteria that are available to every native speaker, literate or not.

> Suppose you came as an explorer into an unknown country with a language quite strange to you. In what circumstances would you say that the people there gave orders, understood them, obeyed them, rebelled against them, and so on?

> The common behavior of mankind is the system of reference by means of which we interpret an unknown language. [**PI** §206]

The criteria for correct usage come from our understanding of the "common behavior of mankind."[23] (NB: It is important to note that Wittgenstein is *not* equating behavior and meaning, only that behavior is the "system of reference" with which we can "interpret an unknown language." See Part II, sections "The Foundation of Language in Instinctive Behavior" and "Why Wittgenstein is not a Behaviorist.")

**Relevance to Information Systems:** Plays are fundamentally stories. So if the best example of a word's use is a scene in a play, then the fundamental structure of linguistic meaning is the narrative.[24] Insofar as language is used in information systems to represent or describe what is stored on the system, its meaning, too, may be more faithfully

---

[22] This lack of a set of "central criteria" for correct meaning comes out in Hilary Putnam's notion of the "Division of Linguistic Labor." Putnam insists that the meaning of a word is not one thing that can be in any single person's possession, but is distributed among a variety of speakers who use the word in question for a variety of activities and purposes. [**Representation and Reality**, p. 22ff. MIT Press, Cambridge, MA, 1996.]

[23] The philosopher W.O. Quine wrote in his famous essay on "the indeterminacy of translation," that if someone speaking a language we did not understand, consistently uttered the word "gavagai" in the presence of a rabbit we still could not be assured that "gavagai" actually meant "rabbit"—it could mean, he claimed, "undetached parts of rabbits" or "rabbithood" [**Word and Object**, §12, p. 52. MIT Press, 1960]. Wittgenstein admitted that language is extremely variable and we may have trouble understanding a language we don't know. But there is more of a common basis for mutual understanding than Quine would admit to. In so far as the activities and practices are the same between two linguistic groups, there will be mutual understanding of their respective languages (Quine makes no claim that the activities of the speakers are at all dissimilar). Differences of meaning are proportional to the differences in respective activities and practices. Since the practice of pointing out and naming things like rabbits is common to virtually all linguistic groups, it is doubtful that "gavagai," when uttered in the presence of a rabbit scurrying by, would ever mean "undetached rabbit parts" in any major languages. In short, the word "rabbit" and the practice of pointing things out, have roles in our activities, but the description "undetached rabbit parts" does not (at least for activities other than philosophical ones).

[24] Mark Turner takes the narrative to be even more basic. For him, the narrative is the fundamental structure of not just language, but cognition itself. He distinguishes between the narrative and the parable. The parable is

interpreted in terms of a narrative. This fits more closely with the idea that information is most productively seen as part of an activity or practice—an activity or practice, like a story or a play, can be described as a sequence of events over time.[25]

> 3. When I think in language, there aren't "meanings" going through my mind in addition to the verbal expressions: the language is itself the vehicle of thought. [**PI** §329]

Language, that is, speech or writing, is not a *product* of thought, but a *means* by which we think. This reversal of the usual way of looking at the relation between what we say or write and what we think, is one of Wittgenstein's most important insights. Language gives us words, phrases and ways of expressing ourselves that serve as a set of implements with which we carry out the activities of speaking, understanding others, and thinking. In his primary later work, **Philosophical Investigations**, Wittgenstein's discussion of the nature of language is replete with metaphors of tools and implements (see Part II of this manuscript: "Words are Like Tools and Language Use is Like Tool Use").

If language is a vehicle of thought, then a number of important consequences follow. In the first place, the words and expressions that we have available in our language represent a kind of limit not only for what we say, but, more importantly, for what we can think, verbally. As the painter is limited by the kinds of paint she has, and the size and style of her brushes, so too are we limited in our verbal thought by the kinds of words and expressions we have available to us—the words and expressions we understand and know how to use. Language is not a straitjacket, though, we can use common words and expressions in new and creative ways, of course. But our verbal thoughts are quite clearly *anchored* in our language, and though we may find ways to express new ideas, what we say must be grounded in the bedrock of our common tongue.

One of the ways in which we use familiar expressions to express new ideas is through metaphors. When Wittgenstein tells us that much of our language use is like a "game" he gives us a clearer sense of the intimacy, dynamics and complexity of language use than he could by describing the detailed processes of expression without the analogy to games [see the section "Language Games" in Part II]. Wittgenstein's frequent use of metaphors in his writings enabled him to stretch the boundaries of philosophical expression into the new areas he wanted to discuss [see the section "Philosophy of Language and Metaphor" in Part II].

Of course, we can also think with images and sound, and Wittgenstein was not denying that this kind of thought occurs, too, but he was primarily concerned with the conduct of philosophy, and philosophy must be written or spoken.

**Relevance to Information Systems:** If the words of language represent a toolset for expression, and, as such, define the limits of our ability to think and express ourselves

---

a narrative that the listener has interpreted in a way that makes it meaningful for him. [**The Literary Mind**. Oxford University Press, Oxford, 1996.]

[25] The idea that information is best seen as part of an activity or practice, was one of the central themes of Blair's **Language and Representation in Information Retrieval**. Elsevier Science, New York, 1990.

verbally, then the words used to represent information must define the limits of our ability to think about and to express our information needs.

> 4. Our craving for generality has another main source: our preoccupation with the method of science. I mean the method of reducing the explanation of natural phenomena to the smallest possible number of primitive natural laws .... Philosophers constantly see the method of science before their eyes, and are irresistibly tempted to ask and answer questions in the way science does. This tendency is the real source of metaphysics, and leads the philosopher into complete darkness. I want to say that it can never be our job to reduce anything to anything, or to explain anything. Philosophy really is "purely descriptive!" [**BB** p. 18]

Although Wittgenstein made this statement against reduction in the **Blue and Brown Books**, one of his earlier writings, his distrust of reduction was an attitude that he carried with him through the remainder of his work.[26] For Wittgenstein, we cannot reduce linguistic meaning to anything more primitive than what we say and do. It is true that we ask for and give definitions of individual words, but asking for and giving definitions is a particular kind of language game, and we can only understand the definition of an individual word if we already understand its general role in our language. Further, these individual definitions of words are not the complete meanings of words, nor can they usually be put together to arrive at the meaning of sentences or longer text—that is, the meaning of a sentence is not always *reducible* to some aggregate of the meanings of the individual words it contains. If we want to understand the meaning of a sentence we must look at how it is used—this is the most basic level of analysis that we can do in language[27] [This point will be presented in more detail in a subsequent section discussing the Augustinean model of language.].

**Relevance to Information Systems:** Since we cannot generally reduce ordinary language to more primitive components of meaning without losing some of the meaning that emerges from its use, we should not expect that a statement that requests information of some kind could be analyzed into more primitive components without some semantic deficit, either. Like language, if we want to understand the meaning of an information request we need to look at how the expression is used, that is, to look at the background of the person making the request, the purpose or rationale for his/her request, the activity that the request serves, and the particular circumstances in which the request was made. Each of these can have a bearing on determining what the request "means," that is, what information would prove useful to the inquirer. Wittgenstein was quite clear in his insistence that

---

[26] Strictly speaking, the **Blue and Brown Books** is not one of Wittgenstein's "writings" in the sense that **TLP** and **PI** are. It is actually a collection of class notes put together by some of his students in the late 1930's as material to accompany his seminars. Although they were not written entirely by Wittgenstein, they do reflect the issues he was grappling with in his seminars, and the notes themselves were thoroughly vetted by Wittgenstein before being distributed. Further, the ideas expressed in **BB** are entirely consistent with the issues that he worked on during the remainder of his life. **BB** is often referred to as "Preliminary Studies for the **Philosophical Investigations**"

[27] Wittgenstein would have a particular disagreement with the kinds of reductive analysis of language that Chomsky engaged in with his "Transformational Grammar." Chomsky's linking of the "surface structures" of language with "deep structures" in the mind, commits two major mistakes on Wittgenstein's analysis: first, in designating "deep structures" as the foundation of language it makes internal, unconscious mental phenomena the determinant of meaning; and, second, it assumes that the expressions we use can be transformed, or reduced, into more primitive units which form the "building blocks" of expression.

we can best express ourselves in ordinary language. This means that our best means of articulating what we want from an information system is with our ordinary, everyday language. When we "reduce" a searcher's information request, stated in ordinary language, to a set of search terms, some loss or distortion of the searcher's meaning must inevitably result.

> 5. Our language can be seen as an ancient city: a maze of little streets and squares, of old and new houses, and of houses with additions from various periods; and this surrounded by a multitude of new boroughs with straight, regular streets and uniform houses. [**PI** §18]

One of Wittgenstein's most compelling metaphors is of language being like a *city*. We live in our language in the same way that we live in our cities. We find our way about our cities by *doing* things, that is, by engaging in our day-to-day activities. It is through the conduct of such day-to-day activities that we learn our way about a city, and it is through the conduct of day-to-day activities that we learn our way about our language. It is also the case that a large city offers many alternative routes for going from one place to another. Similarly, language offers us many alternative ways to say the same thing. The city, no matter how large, can also exist quite efficiently without any kind of central planner or controller. The day-to-day local interactions within cities even as large as New York are sufficient to keep the whole functioning, and, although the city, at any time, contains only a limited supply of essentials such as food, it never runs out, even during major disruptive events—snow storms, labor strikes, power outages, etc. Likewise, language needs no central authority to control usage, it needs only the day-to-day interactions of its native speakers to establish and retain its meaning [See the section "Language as a City" in Part II for a more detailed discussion of this metaphor.].

**Relevance to Information Systems:** The language of information systems is like a city, too, in the following ways: like a city, it is constantly evolving and changing in response to the activities that it serves; like a city, there is no need for central planning if the day-to-day interactions of searchers can provide feedback about how language is used, that is, if searchers can learn the correct usages of search terms from day-to-day searches on the system. The meaning of search terms or descriptions of information *emerges* from the day-to-day interactions of users.

> 6. But how many kinds of sentences are there? Say assertion, question, and command?—There are countless kinds; countless different kinds of use of what we call "symbols" "words," "sentences." And this multiplicity is not something fixed, given once for all; but new types of language, new language-games, as we may say, come into existence, and others become obsolete and get forgotten.[28] [**PI** §23]

---

[28] Wittgenstein's quotation continues:

"Here the term 'language-game' is meant to bring into prominence the fact that the speaking of language is part of an activity, or of a form of life.

Review the multiplicity of language-games in the following examples, and in others:

    Giving orders, and obeying them—
    Describing the appearance of an object, or giving its measurements—

Language does not work in just one way. Wittgenstein begins **PI** with a quotation from the medieval philosopher Augustine who believed that language worked in just one way: words stood for "objects," and language was taught by pointing to the objects that words "stood for." Subsequent theorists have assumed a more subtle version of Augustine's view holding that words stand for "ideas" (Locke) and that the meaning of a word must be somehow "pointed out" or explained to the person who wants to understand it. But for Wittgenstein, language is so diverse that there are uncountable different ways in which it can be used—language works in different ways in different activities or practices. Expressions can be used to "point things out," but they can also be used ironically, sarcastically, or metaphorically. Language is not used primarily to assert facts, as Frege, Russell and the early Wittgenstein, in **TLP**, believed, but can be used to make a promise, tell a joke, order someone to do something, to lie, to exaggerate, to collude, to elaborate, to tell a story, declare war, or to do any of an uncountable number of things.[29] Each of these uses involves using words or expressions in different ways, and fits in to the needs of many kinds of activities or practices [Wittgenstein's critique of the Augusintean view of language will be presented in more detail in Part III].

**Relevance to Information Systems:** The words used in information systems to represent information are not just a collection of "labels" that are somehow linked to information content, like Augustine's notion that words were names for objects. In information systems, like language in general, language is not used in just one way. An index term can describe the intellectual content of information, like a subject description, but it can also link the information to activities or practices. It can be used to assert the quality of the information, to link information to other related information, to describe how the information has been used (as a contract, a directive, a declaration, etc.), or it can name various contextual information such as the author(s) of the text, the date it was published, the source of the

Constructing an object from a description (a drawing)—
Reporting an event—
Speculating about an event—
Forming and testing a hypothesis—
Presenting the results of an experiment in tables and diagrams—
Making up a story; and reading it—
Play-acting—
Singing catches—
Guessing riddles—
Making a joke; telling it—
Solving a problem in practical arithmetic—
Translating from one language into another—
Asking, thanking, cursing, greeting, praying. [**PI** §23]

---

[29] This assertion that language works in a multitude of ways runs counter to Searle's theory of "speech acts." He takes issue specifically with Wittgenstein's claim, in paragraph 23 of **PI** (see previous footnote) that there are "countless different kinds" of "symbols, words, sentences" [J.R. Searle. **Speech Acts, An Essay in the Philosophy of Language**. Cambridge University Press, NY, 1969]. But Searle's speech acts are based on John Austin's "illocutionary acts," and Austin is quite clear that there are many more kinds of illocutionary acts than the few he names. Austin, like Wittgenstein, insisted on the great multiplicity of usage types—that is, illocutionary verbs—in natural language, though he limited their number to between 1,000 and 10,000 [Austin. **How to Do Things With Words**, p. 150, n. 1. Oxford, 1962.]. Searle, in contrast, limits the number of different kinds of Speech Acts to fewer than ten.

text (a magazine, journal, book, etc.) or the type of document it is (published article, report, minutes of a meeting, evaluation, white paper, regulation, etc.). The number of ways that index terms can be used is similar to the number of ways that the information represented by the terms can be used.

> 7. Many words . . . then don't have a strict meaning. But this is not a defect. To think it is would be like saying that the light of my reading lamp is no real light at all because it has no sharp boundary. [**BB** p. 27]

One of the clearest breaks that Wittgenstein had with his "old way of thinking" in the **Tractatus Logico-Philosophicus** (**TLP**) was over the "determinacy of sense." The early Wittgenstein, along with his mentors Frege and Russell, believed that for language to be useful in philosophical analysis it must have a "strict determinacy of sense," that is, each word must have a precise, unambiguous "sense," or meaning, that was independent of context, and would hold for all its possible uses. In the **Tractatus** Wittgenstein was even more insistent on the importance of the "determinacy of sense" in language—he felt that "sense" in language had to be determinate for language to be possible at all:

> The requirement that simple signs be possible is the requirement that sense be determinate. [**TLP** §3.23]

This is a natural consequence of the belief that language was used primarily to assert facts or make propositions. Facts are generally precise and unambiguous and stay this way regardless of the context in which they appear. So, if language asserts facts, then it should be as precise and unambiguous as the facts it represents. Ordinary language, of course, is not like this, as Frege and Russell saw.[30] Consequently, they believed, ordinary language could not be used for the precise kinds of philosophical analysis—the clarification of propositions, or statements of fact—that Frege and Russell wanted to systematize. This was the reason for their insistence on constructing a language of philosophical analysis that did not have the indeterminacy of ordinary language. The exemplar of the kind of language they wanted was formal logic. In logic, as in other formal systems such as arithmetic, the symbols *do* have a strict determinacy of sense—a logical or mathematical symbol, once defined, means the same thing regardless of what equation or "phrase" it is used in. This is why logic was the model for the philosophical language Frege and Russell deemed necessary for analysis. In his early work, Wittgenstein believed, too, that ordinary language lacked the necessary strict determinacy of sense for philosophical analysis, but he differed from Frege and Russell about how to mitigate this indeterminacy of language. In the **TLP**, Wittgenstein argued that we didn't need another language to do philosophy, we just needed to be more careful about our use of ordinary language. In short, ordinary language clearly makes sense most of the time, so it must have an underlying logical order. The ambiguity of language is an illusion, Wittgenstein asserted, we just need to bring out this underlying logical order, the strict determinacy of sense, of language. What Wittgenstein saw later was that language did *not* get its determinacy from some underlying logical order. Like logic, its "crystalline purity" was "not a result of investigation, it was a requirement." Wittgenstein grew to see that "Many words. . . then don't have a strict meaning. But this is not a defect."

---

[30] As Russell wrote in the his introduction to Wittgenstein's **TLP**: "In practice, language is always more or less vague, so that what we assert is never quite precise." [**TLP**, p. x]

(*supra*) Language can be as determinate as necessary. To see the indeterminacy as a defect is to look at ordinary language as if it were a kind of formal calculus, which it is not. We can make language very precise if we want, not by bringing out some kind of hidden logical underpinning, but by looking at the context, circumstances and practices in which language is used. In **Philosophical Investigations**, Wittgenstein takes a final jab at Frege's insistence on a strict determinacy of sense in language:

> Frege compares a concept to an area and says that an area with vague boundaries cannot be called an area at all. This presumably means that we cannot do anything with it.—But is it senseless to say: "Stand roughly there?" Suppose that I were standing with someone in a city square and said that. As I say it I do not draw any kind of boundary, but perhaps point with my hand—as if I were indicating a particular spot. [**PI** §71]

Wittgenstein's analogy between the meaning of words and the boundary of the light from his reading lamp (*supra*) is particularly *apropos*. The determinacy of sense, or precision of meaning, needs only to be as strict as is necessary for the task at hand—there is no absolute level of determinacy to which all language aspires. In the same way, a light doesn't have a single standard of brightness, it only needs to be bright enough for the task at hand. To decry the ambiguity of individual word meaning is to apply the requirements of a particular Language Game, scientific discourse, for example, to all language. This, of course, is exactly what Russell did, since he claimed that the purpose of philosophy was to clarify scientific assertions about facts.

**Relevance to Information Systems**. As we will see in Part III, the "determinacy of sense" is one of the central issues of Information Systems. In fact, we can line up different kinds of information systems along a spectrum of their respective determinacies of representation. At one extreme we have data base management systems which provide access to highly determinate information such as names, addresses, phone numbers, and account balances. Such data items are unambiguous and stay pretty much the same regardless of the context in which they appear. At the other end of the determinacy spectrum would be a document retrieval system which provides access to the "intellectual content" of a large collection of documents many of which deal with similar topics. Here, a retrieval request for a document detailing the "reasons for the failure of the Marxist economic model" might be very hard to make without retrieving too many or too few documents (imagine devising a search query to do this for use with an Internet Search Engine). This issue of the determinacy of sense in Information Systems will be central to the discussion of Part III.

> 8. We want to establish an order in our knowledge of the use of language: an order with a particular end in view; one of many possible orders; not *the* order. [**PI** §132]

Although Wittgenstein was concerned with identifying errors in our use of language, his goal was to correct philosophical mistakes that arose because of errors in language, not to correct errors in ordinary usage or to create a separate, more precise, ideal language. Although Wittgenstein felt that philosophers needed to be reminded of their "diseases of thinking" which arise from the misuse of language, he believed that ordinary native speakers of a language needed no such reminders.

**Relevance to Information Systems:** Similarly to ordinary language, the language of representation in information systems does not aspire, in theory or in practice, to some ideal language. As with ordinary language, the principal criterion of quality for a search language is whether or not its users, that is, searchers, can *use it* to find what they want. It is also the case that, like ordinary language, the language of information representation and searching is not a fixed entity, but changes insofar as the uses for the information that it provides access to change. Further, just as Wittgenstein insisted that ordinary language usage is the final arbiter in questions of meaning, even philosophical meaning, ordinary language usage will likewise be the final arbiter in questions of meaning about search requests and information representations. Specifically, ordinary language is the best medium for us to express our information needs, and any subset of ordinary language that may be used as an access language to an information system will be correspondingly *less effective* than ordinary language for searching. This poses a particular problem in information systems, namely, that if ordinary language is the best medium in which to express our information requirements, then *computerized* information systems, with which we *cannot* interact using our ordinary means of expression, will constrain our ability to express our information requirements.[31] It may mean also that the best content-retrieval systems, in general, are those which can understand the subtle meanings and nuances of information requests stated in ordinary language. At this point, only experienced search intermediaries can do this.

> 9. My method is not to sunder the hard from the soft, but to see the hardness of the soft. [**NB** p. 44]

Ordinary language, and the practices and activities of which they are a part are our primary references for meaning, even philosophical meaning. The "softness" of meaning in ordinary language cannot be improved by the "hardness" of logic. We don't need a separate more determinate language to eliminate the indeterminacies in meaning, we just need to be more careful about how we *use* language. Further, no such more determinate language is possible since there can be no single comprehensive notion of the determinacy of sense applicable to all uses of language in all contexts and circumstances.[32] The determinacy of sense is not a *property* of individual words, but a function of how those words are used, and can be influenced by the context and circumstances in which they are used or the backgrounds and experience of the participants. The same word can have different levels of determinacy in different usages. More importantly, Wittgenstein saw that when issues of meaning in language come up, as they did frequently in philosophy, the final arbiter of linguistic meaning is not philosophical analysis, but *ordinary usage*.[33] In fact, many of the problems of analytic philosophy arose, Wittgenstein believed, because

---

[31] While there has been work in the fields of Computer Science and Artificial Intelligence to build natural-language front ends to information systems (most commonly, to data base systems where the language, as we have said, is more precise), none of these operates in a fully conversational mode that would be comparable to the kind of semantically rich and complex conversation that an inquirer might have with a reference librarian in a research library.

[32] The notion of the determinacy of sense and its problems is the focus of a major portion of **PI** beginning with §65.

[33] "When I talk about language (words, sentences, etc.) I must speak the language of every day. Is this language somehow too coarse and material for what we want to say? Then how is another one to be constructed?—And how strange that we should be able to do anything at all with the one we have!" [**PI** §120]. Wittgenstein's question is, of course, rhetorical.

philosophers forgot these quotidian constituents of linguistic meaning. Language that has lost its connection to ordinary usage can indeed "bewitch" us.[34] Language leads us into some of these difficulties so frequently and predictably, that Wittgenstein called these systematic errors "diseases of thinking."[35]

In particular, Wittgenstein disagreed with Frege and Russell's belief that, in general, logic would form a better, more precise, foundation for linguistic expression—that language somehow aspired to the "crytaline purity of logic." Wittgenstein came to see that the "rough ground"[36] of ordinary language, as it is caught up in the "hurly-burly"[37] of day-to-day usage, gives us more precision in meaning than philosophers had previously thought possible. This is not a despairing observation, in fact, Wittgenstein reminded his readers that it is surprising how *well* ordinary language works:

> We are under the illusion that what is peculiar, profound, essential, in our investigation, resides in its trying to grasp the incomparable essence of language. That is, the order existing between the concepts of proposition, word, proof, truth experience, and so on. This order is a super-order between—so to speak—super-concepts. Whereas, of course, if the words "language," "experience," "world," have a use, it must be as humble a one as that of the words "table," "lamp," "door."

> On the one hand it is clear that every sentence in our language "is in order as it is." That is to say, we are not striving after an ideal, as if our ordinary vague sentences had not yet got a quite unexceptionable sense, and a perfect language awaited construction by us.—On the other hand it seems clear that where there is sense there must be perfect order.—So there must be perfect order even in the vaguest sentence. [**PI** §§97–98]

**Relevance to Information Systems:** The language of an information system is often created by system designers as they build the system, in much the same way that logicians such as Frege and Russell proposed constructing logical languages to be used in philosophical analysis. Like them, system designers often despair that ordinary language is not precise enough to be useful as a language for content searching, especially on large text or image retrieval systems. But while the language of an information system may be a subset of ordinary language, and its uses may differ, the system designer should not overlook the fact that there may be more precision in ordinary language than he may at first think. In particular, there are often contextual aspects of language that can be used to provide better access to available information. One of the great dangers of building *computerized* information systems is that the computer seems to give us a precision in

---

[34] "Philosophy is a battle against the bewitchment of our intelligence by means of language." [**PI** §109]

[35] "There is a kind of general disease of thinking which always looks for (and finds) what would be called a mental state from which all our acts spring as from a reservoir." [**BB** p. 143]

[36] "The more narrowly we examine language, the sharper becomes the conflict between it and our requirement. (For the crystaline purity of logic was, of course, not a result of investigation; it was a requirement.) The conflict becomes intolerable; the requirement is now in danger of becoming empty.—We have got onto slippery ice where there is no friction and so in a certain sense the conditions are ideal, but also, just because of that, we are unable to walk. We want to walk; so we need friction. Back to the rough ground!" [**PI** §107]

[37] "How could human behavior be described? Surely only by showing the action of a variety of humans, as they are all mixed up together. Not what one person is doing now, but the whole hurly-burly, is the background against which we see an action, and it determines our judgement, our concepts, and our reactions." [**RPP II** §629]

access that is not possible in non-computerized systems. As Wittgenstein remarked about logic, "the crystalline purity of logic was, of course, not a result of investigation; it was a requirement." Some of this belief that the precision of a computer confers an advantage in searching is due to a confusion between *physical access* and *intellectual access. Physical access* consists of the means by which a computer locates and retrieves information whose precise address is known, while *intellectual access* consists of the means by which certain information is determined to be what a particular inquirer will be satisfied with, having submitted a specific query to the information system. Computers *can* give us rapid, precise *physical access*, no matter how complex the search requirements are, as long as they are decidable. But rapid, precise *physical access* does not always guarantee rapid, precise *intellectual access,* because in some situations, such as with text or image retrieval, the specific item(s) of information that will satisfy an inquirer often remain(s) an open question during the search. Of course, in data retrieval, where the descriptions of available information—addresses, phone numbers, account balances, etc.—are quite precise, then improvements in *physical access* can improve *intellectual access*. Where does this misconception about precision in language come from? For Frege and Russell, formal logic was the model for all language, leading them to believe that the precision of logical propositions could be attained for any semantic meaning. In the same way, information systems designers have often assumed that the precision of representation and retrieval of a data retrieval system—one that provides access to, for example, names, addresses and phone numbers—is attainable in text or image retrieval systems where the searchers are looking for items with specific intellectual content. But it is clear that, as Wittgenstein states, ". . . every sentence in our language 'is in order as it is' " (*supra*). If there is no better way to express ourselves than in ordinary language, then it is also evident that the farther away from ordinary usage that we get, the less likely we are to express ourselves well. Insofar as information systems employ means of expression different than our ordinary means of expression, our ability to express our information needs will be impoverished. Currently, the only "information system" that uses the full range of expression of our ordinary language is another human being.

> 10. One of the most dangerous of ideas for a philosopher is, oddly enough, that we think with our heads or in our heads.
>
> The idea of thinking as a process in the head, in a completely enclosed space, gives him something occult. [**Z** §§605–606]

Thinking and meaning are not entities that have simple "locations" in the way that the anatomical parts of the brain have physical places. Our thinking is frequently assisted by implements that exist outside of our skulls: Much information that we use does not exist "in our heads" but in books, audio tapes, computer data bases, the notes we make to ourselves and the people with whom we interact. In some instances, our ability to think may require a calculator, a computer spread-sheet, or just pencil and paper. This sort of augmentation of the human intellect has been called "scaffolding," and will be discussed in Part III.

If our thought processes must sometimes be assisted by implements and information storage media that are *external* to our heads, then there can be no comprehensive