# Can We Ever Escape From Data Overload?
# A Cognitive Systems Diagnosis

**David D. Woods**
Cognitive Systems Engineering Laboratory
Institute for Ergonomics
The Ohio State University
210 Baker Systems, 1971 Neil Ave.
Columbus, OH 43210 U.S.A.
woods.2@osu.edu

**Emily S. Patterson**
Cognitive Systems Engineering Laboratory
Institute for Ergonomics
The Ohio State University
210 Baker Systems, 1971 Neil Ave.
Columbus, OH 43210 U.S.A.
patterson.150@osu.edu

**Emilie M. Roth**
Roth Cognitive Engineering, Inc.
89 Rawson Rd.
Brookline, MA 02445-4509 U.S.A.
emroth@mindspring.com

*Cognition, Technology and Work*, in press

## ABSTRACT

Data overload is a generic and tremendously difficult problem that has only grown with each new wave of technological capabilities. As a generic and persistent problem, three observations are in need of explanation: Why is data overload so difficult to address? Why has each wave of technology exacerbated, rather than resolved, data overload? How are people, as adaptive responsible agents in context, able to cope with the challenge of data overload?

In this paper, first we examine three different characterizations that have been offered to capture the nature of the data overload problem and how they lead to different proposed solutions. As a result, we propose that (a) data overload is difficult because of the context sensitivity problem – meaning lies, not in data, but in relationships of data to interests and expectations and (b) new waves of technology exacerbate data overload when they ignore or try to finesse context sensitivity. The paper then summarizes the mechanisms of human perception and cognition that enable people to focus on the relevant subset of the available data despite the fact that what is interesting depends on context.

By focusing attention on the root issues that make data overload a difficult problem and on people's fundamental competence, we have identified a set of constraints that all potential solutions must meet. Notably among these constraints is the idea that organization precedes selectivity. These constraints point toward regions of the solution space that have been little explored. In order to place data in context, designers need to display data in a conceptual space that depicts the relationships, events, and contrasts that are informative in a field of practice.

KEYWORDS: agent, alarm, context, data overload, information visualization, workload

## DATA OVERLOAD IS A GENERIC, DIFFICULT PROBLEM

Information is not a scarce resource. Attention is.
Herbert Simon[1]

Each round of technical advances, whether in artificial intelligence, computer graphics, or electronic connectivity promises to help people better understand and manage a whole host of activities from financial analysis to monitoring data from space missions to controlling the national air space. Certainly, this ubiquitous computerization of the modern world has tremendously advanced our ability to collect, transmit and transform data producing unprecedented levels of access to data.

However, our ability to interpret this avalanche of data, i.e., to extract meaning from artificial fields of data, has expanded much more slowly, if at all. In studies across multiple settings, we find that practitioners are bombarded with computer processed data, especially when anomalies occur. We find users lost in massive networks of computer based displays, options and modes. For example, one can find a version of the following statement in most accident investigation reports:

> "although all of the necessary data was physically available, it was not operationally effective. No one could assemble the separate bits of data to see what was going on" (Joyce and Lapinski, 1983).

The challenge has become finding what is informative given our interests and needs in a very large field of available data.

The paper is organized as follows. To set the stage, we characterize how technology change has created a paradoxical situation and, we introduce people as model of competence through a historical example. From this base we summarize the three different major characterizations of the data overload problem. We then provide a "diagnosis" of what makes data overload a difficult problem based on a synthesis of results from past studies that examine how new computerized devices can help overcome or can exacerbate data overload related problems in control centers such as mission control for space shuttle operations, highly automated aviation flight decks, computerized emergency operations control centers in nuclear power plants, and surgical anesthetic management systems in operating rooms. Given this background, we can see how the typical solutions to the data overload problem avoid confronting the heart of the matter directly, remaining

---

[11] In written publications, Simon has made this point several times:
"The information-processing systems of our contemporary world swim in an exceedingly rich soup of information, of symbols. In a world of this kind, the scarce resource is not information; it is the processing capacity to attend to information. Attention is the chief bottleneck in organizational activity ..." (Simon, 1976, p. 294).
"A design representation suitable to a world in which the scarce factor is information may be exactly the wrong one for a world in which the scarce factor is attention." (Simon, 1981, p. 167).

content to nibble away at the edges through indirect means. Finally, we outline a direction for progress towards more effective solutions to data overload relying on people as a competence model.

**The Data Availability Paradox**

Our situation seems paradoxical: more and more data is available in principle, but our ability to interpret what is available has not increased. On one hand, all participants in a field of activity recognize that having greater access to data is a benefit in principle. On the other hand, these same participants recognize how the flood of available data challenges their ability to find what is informative or meaningful for their goals and tasks (Miller, 1960). We will refer to this as the *data availability paradox*. Data availability is paradoxical because of the simultaneous juxtaposition of our success and our vulnerability. Technological change grows our ability to make data readily and more directly accessible – the success, and, at the same time and for the same reasons, the change increasingly and dramatically challenges our ability to make sense of the data available – the vulnerability.

**"A Little More Technology Will Be Enough"**

> Criando difficuldades para vender facilidades
> (creating difficulties to sell solutions).
> common Brazilian saying

As the powers of technology explode around us, developers imagine potential benefits and charge ahead in pursuit of the next technological advance. The claim is that data overload and other problems will be solved by significant advances in machine 'information' processing, i.e., the technology for creating sophisticated graphics, for connecting distant people together, and for creating intelligent software agents.

However, after each round of development, field researchers continue to observe beleaguered practitioners actively trying to cope with data overload in one form or another. This is a fundamental finding, repeatedly noted in many fields of practice and with many kinds of technology (e.g., Woods, 1995a; Woods and Patterson, 2000). When viewed in context, systems, developed putatively to aid users, often turn out to create new workload burdens when practitioners are busiest, new attentional demands when practitioners are plagued by multiple channels/voices competing for their attention, and new sources of data when practitioners are overwhelmed by too many channels spewing out too much "raw" data (Woods et al., 1994, chapter 5).

In practice, new rounds of technology development become yet another voice in the data cacophony around us. Ironically, the major impact has been to expand the problem beyond specialized technical fields of activity (an airplane cockpit or power plant control room) to broader areas of activity (web based activities we engage in everyday).

## People Are Competent to Find the Significance of Data: The "Wow!" Signal

The irony of the data availability paradox is that people in general are very good at finding the significance of data under many conditions. For example, Figure 1 is a printout of numbers and letters in a structure of columns and rows. An observer highlighted some of these data elements, writing the note "**Wow!**" in the margin. Clearly, these data elements were highly significant to this observer. This is how the participants later described the Wow! Signal:

> "Are we alone or are there other beings out there across the immense reaches of space who might be sending out radio signals we could hear?" Radio observatories, such as the Ohio State - Ohio Wesleyan radio observatory, try to answer this question by searching for signals which might indicate an extraterrestrial intelligent origin.
>
> "In mid-August (1977) Jerry Ehman showed Bob Dixon, Dick Arnold and me a section of new computer print-out with all of the characteristics that one might expect from an extraterrestrial beacon signal. Jerry's amazement was reflected by the words "Wow!" which he had written on the margin of the print-out (Figure 1). Bob, Jerry, Dick and I had urgent discussions about its significance. We soon were referring to it as the "Wow!" signal.
>
> The print-out format, which Bob Dixon had designed, consisted of 50 columns, one for each channel, with a single digit printed every 12 seconds indicating the signal level in that channel in units above the background level (the technical term for the unit used is one "standard deviation" or one "sigma"). A blank signified that the level was at zero. Any number above 4 or 5 might be considered as significant and probably not due to some random fluctuation. In order to accommodate levels above 9 with a single character, Bob arranged that the computer run through the alphabet with A for 10 through Z for 35.
>
> What Jerry had noted was a sequence of characters in Channel 2 running: 6, E, Q, U, J, 5. When plotted up they produced a pattern which matched exactly (within measurement error) the telescope antenna pattern. This told us that the source was very probably celestial, that is, fixed with respect to the star background and that it passed through the telescope beam with the earth's rotation. It was strong (30 sigmas or 30 times the background) and because it appeared in only one channel it was narrow-band (width 10 kilohertz or less). But even more significant, it was intermittent. A steady signal would have appeared two times on the record a few minutes apart as our telescope with its twin-beam scanned the sky. (The possibility that only one horn was functioning at the time can be ruled out because the two horns are balanced and, if one were out, the system would have been inoperative.) So it was an "on and off" signal! Was it intended for us? We decided that we should continue to scan the same region of sky on the chance that the signal might reappear. But it never did, and after weeks of patient listening we moved on with our survey to other parts of the sky. Kraus, J., 1979.

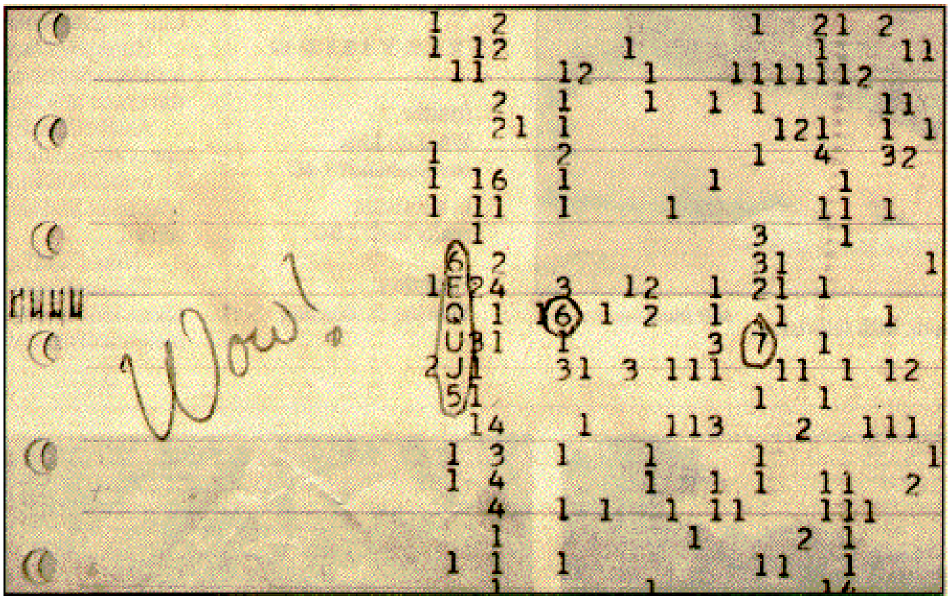


**Figure 1. The "Wow!" signal.**

Several things should strike us as we consider this example. To us, the data elements look like a meaningless mass of numbers and letters, since we

- lack the knowledge of this observer (a radio astronomer),
- lack any knowledge about what and how the elements symbolize (e.g., they represent radio telescope signals coded as the number of standard deviation units above background level),
- have no particular expectations about what is background, typical, or recent (the norm for years has been random, low level signals)
- do not know the goals of the observer (searching space for patterns of signals that might indicate an extraterrestrial intelligent origin),
- do not know how patterns of signals that might indicate an extraterrestrial intelligent origin would be expressed in the representation.

But the data elements are not meaningless for the experienced, knowledgeable observer. While the representation looks quite crude, it does provide some support. The data is selected, pre-processed, and organized to enable experienced observers to scan for patterns. The observers are looking for an unknown, new signal, yet they can determine some properties or relationships to look for -- departures from background, patterns associated with signals coming from different kinds of sources and, particularly, sources of extraterrestrial intelligent origin. The data is laid out in parallel. Given their knowledge in the field of practice and experience at scanning this representation, observers can recognize interesting patterns that stand out against the background. For example, radio astronomers at one point noticed an unusual pattern which further investigation revealed as a new natural phenomenon -- pulsars.

While knowledgeable, experienced observers can find significant patterns in this data field, we are also struck by the fragility of this process given representations and tools like this printout. First, to succeed at all requires great investment in human expertise – people knowledgeable in the field of practice and practiced at observing through this representation. Second, even though people can succeed, we often find cases where the people involved miss significant aspects of the data field. Third, the kinds of representations developed in this case and others (such as status boards, annunciator panels, logs, and trend plots in traditional control rooms) are technologically crude. It seems obvious that applying more sophisticated computer processing and graphics capabilities should lead to more effective representations and tools with respect to data overload.

This example illustrates that people can find the significance in a field of data, i.e., people possess the competence to find the significance in a field of data though they may not always exhibit this competence in practice for specific cases. The questions for us to consider are: How are people able to do this at all? How does computer technology affect people's ability to do this? How can we design visualizations to help people do this?

### What Needs to be Explained in a Diagnosis of Data Overload?

As a generic and persistent problem, three observations are in need of explanation:
1. Why is data overload so difficult to address?
2. Why has each wave of technology exacerbated, rather than resolved, data overload?
3. How are people, as adaptive responsible agents in context, able to cope with the challenge of data overload?

Developing explanations for these phenomena is what we mean by a diagnosis of data overload. Such a diagnosis is important because, ultimately, the goal is to innovate new approaches that will enable our escape from the paradox of data availability.

## CHARACTERIZATIONS OF DATA OVERLOAD

There are three basic ways that the data overload problem has been characterized:
1. As a clutter problem where there is *too much data* on the screen: therefore, we can solve data overload by reducing the number of data units that are displayed,
2. As a *workload bottleneck* where there is too much to do in the time available: therefore, we can solve data overload by using automation and other technologies to perform activities for the user or to cooperate with the user during these activities, and

3. As a problem in *finding the significance of data* when it is not known a priori what data from a large data field will be informative: therefore, we can solve data overload by representing the data field in a way such that the significant data naturally emerges from the virtual perceptual field.

## Clutter

> "Clutter and confusion are failures of design, not attributes of information."
>
> Tufte, 1990, p. 51

The first way that people have characterized data overload is simply that there is "too much stuff." This problem formulation led designers to try to reduce the available data. This approach arose in the early 1980's as a "solution" to the problem of "clutter" in the design of individual displays. The approach led developers to ask: how much is too much for people or what is the maximum rate of data people can process? Developers proposed guidelines for display design that limited the number of pixels that could be lit on the screen (given technological advances this measure of screen density is obsolete, but other ways to define what are too many screen elements can, and have been, proposed).

This has not proven to be a successful or fruitful direction in solving data overload and has faded in large part because:

- it misrepresents the design problem -- see for example Tufte, (1990) and Zhang and Norman (1994); one specific example is that reducing data elements on one display increases people's need to navigate across multiple displays (Woods and Watts, 1997),
- it is based on erroneous assumptions about how human perception and cognition work; for example, the questions about maximum human data processing rates are meaningless and misleading because among other things people re-represent problems, re-distribute cognitive work, and develop new strategies and expertise as they confront clutter and complexity,
- it is incapable of dealing with the context sensitive nature of meaning; in some contexts, some of what is removed will be relevant.

Systems that reduce or filter available data are brittle in the face of context sensitivity. First, some of usually unimportant data may turn out to be critically informative in a particular situation. For example, one nuclear power plant accident scenario is difficult precisely because the critical piece of data is usually unimportant (Roth, Woods, and Pople, 1992). Second, some data which seems minor now may turn out to be important later after new events have changed the context.

Instead of removing data to reduce clutter, some propose to push some data into the background and provide navigation techniques so that users could call up this data should they judge it relevant. However, this does not help the practitioner to recognize what is relevant or help direct the attention of the practitioner *before the practitioner knows where to look or what to look for*.

It is striking that this "solution" to the clutter problem -- reducing the displayed data or hiding it until a user asks for more data -- runs counter to the technological trends. If one of the benefits of certain technologies is increased access to data, it is ironic that people have to throw away some of that access to cope with the complexity of trying to work with the available data.

**Workload Bottleneck**

The second characterization of data overload often appears in settings where access to data has grown quickly and explosively. In these contexts, such as intelligence analysis but also in web based activities, participants use the words "data overload" to means they are experiencing a workload bottleneck -- there are simply too many individual data units to examine them all manually in the time that is available. As a result, one can observe coping strategies and associated failure modes associated with workload bottlenecks in general (Miller, 1960; Hollnagel, Bye and Hoffman, 2000).

Intelligence analysis-like situations are examples where workload is a potentially useful way to think about data overload. Previously, analysts were expected to read the vast majority of the reports that were available to them in order to provide a synthesized assessment and recommendations for action on a topic. With the workload characterization of data overload, analysts now express a need for an "agent" to help them with their activities. For example, machine agents could potentially prioritize or summarize reports for the analyst. Notice that with the workload characterization, solutions no longer are focusing on reducing data at the level of individual data units, but now they are focusing on making a person's cognitive activities more tractable.

An important distinction in aiding approaches to solve the workload bottleneck version of data overload is whether or not the approach requires a strong or weak commitment to the automation being "correct." Brittleness of machine processing, particularly in complex, high-consequence domains, is a serious issue in the design of cognitive systems (e.g., Smith, McCoy, and Layton, 1997). Approaches such as filters, summarizers, and automated search term selectors (e.g., Maes, 1998, Marx and Schmandt, 1996, Brann, Thurman, and Mitchell, 1996) are strongly committed to the machine processing being correct. Methods that are more weakly committed to machine pre-processing include using automation to index, cluster, organize, highlight, sort, and prioritize elements in a data field, (e.g., Oakes and Taylor, 1998, Letsche and Berry, 1997) and "cooperative machine agents" that notify, remind, or critique a human partner (e.g., Gruen et al., 1999, Guerlain et al., 1999, Fischer and Reeves, 1992).

Although the workload characterization is a potentially useful way to think about data overload, the findings clearly show that automation support is necessary but not sufficient to create useful systems. Introducing autonomous machine agents changes the cooperative structure creating new roles, new knowledge requirements, new judgments, new demands for attention, and new coordinative activities. The automation must be observable and directable in order to avoid patterns of coordination breakdowns such as clumsy automation and automation surprises (Sarter, Woods, and Billings, 1997, also see Maes and Schneiderman, 1997 for a debate on interface agents vs. direct manipulation techniques that touch on some of these issues).

**The Significance Of Data**

> It is of the highest importance in the art of detection to be able to recognize, out of a number of facts, which are incidental and which are vital.
>
> Sherlock Holmes/A. Conan Doyle

The starting point for this third sense of data overload is recognizing that large amounts of potentially available data stress one kind of cognitive activity -- focusing in on the relevant or interesting subset of data for the current problem context. When people are unable to assemble or integrate the relevant data, this function in cognitive work has broken down.

People are a competence model for this cognitive activity because people are the only known cognitive system that is able to focus in on interesting material in natural perceptual fields, even though what is interesting depends on context (Woods and Watts, 1997). The ability to orient focal attention to "interesting" parts of the natural perceptual field is a fundamental competency of human perceptual systems (Rabbitt 1984; Wolfe 1992).

> "The ability to look, listen, smell, taste, or feel requires an animal capable of orienting its body so that its eyes, ears, nose, mouth, or hands can be directed toward objects and relevant stimulation from objects. Lack of orientation to the ground or to the medium surrounding one, or to the earth below and the sky above, means inability to direct perceptual exploration in an adequate way (Reed, 1988, p. 227 on Gibson and perceptual exploration in Gibson, 1966)."

Both visual search studies and reading comprehension studies show that people are highly skilled at directing attention to aspects of the perceptual field that are of high potential relevance given the properties of the data field and the expectations and interests of the observer. Reviewing visual search studies, Woods (1984) commented, "When observers scan a visual scene or display, they tend to look at 'informative' areas . . . informativeness, defined as some relation between the viewer and scene, is an important determinant of eye movement patterns" (p. 231, italics in original). Similarly, reviewing reading comprehension studies, Bower and Morrow (1990) wrote, "The principle . . . is that readers direct their attention to places where significant events are likely to occur. The significant events . . . are usually those that facilitate or block the goals and plans of the protagonist."

In the absence of this ability, for example in a newborn, as William James put it over a hundred years ago, "The baby assailed by eye, ear, nose, skin and entrails at once, feels it all as one great blooming, buzzing confusion" (James, 1890, I 488). The explosion in available data and the limits of typical computer-based displays have left us often in the position of that baby -- seeing a "great blooming, buzzing confusion" in the kinds of virtual data fields that technology makes it so easy to create.

What mechanisms underlie the human ability to find the significance of data when acting in natural perceptual fields? Can we use knowledge of these mechanisms to identify constraints and innovate techniques that will help people exhibit this ability when they work in the virtual perceptual fields created by modern technology?

## WHY IS FOCUSING IN ON WHAT IS INTERESTING DIFFICULT? THE PROBLEM OF CONTEXT SENSITIVITY

> Given an enormous amount of stuff, and some task to be done using some of the stuff, what is the *relevant stuff* for the task? (italics in original)
>
> Glymour 1987, p. 65

The cognitive activity of focusing in on the relevant or interesting subset of the available data is a difficult task because what is interesting depends on context. What is informative is context sensitive when the meaning or interpretation of any change (or even the absence of change) is quite sensitive to some but not all the details of the current situation or past situations. Consider this NASA example that describes the response to the discovery of a set of computer alarms linked to the astronauts displays shortly before the Apollo 11 mission:

> A [computer] program alarm could be triggered by trivial problems that could be ignored altogether. Or it could be triggered by problems that called for an immediate abort [of the lunar landing]. How to decide which was which? It wasn't enough to memorize what the program alarm numbers stood for, because even within a single number the alarm might signify many different things. "We wrote ourselves little rules like 'If this alarm happens and it only happens once, don't worry about it. If it happens repeatedly, but other indicators are okay, don't worry about it.'" And of course, if some alarms happen even once, or if other alarms happen repeatedly and

> the other indicators are not okay, then they should get the LEM [lunar module] the hell out of there. (Murray and Cox 1990)

In this example the alarm codes mean different things depending on the context in which they occur. This and other examples reveal that the meaning of a particular piece of data depends on what else is going on, what else could be going on, what has gone on, and what the observer expects or intends to happen.

Formally, information is a relation between the data, the world the data refers to, and the observer's expectations, intentions, and interests (Woods, 1991). Understanding this is critically important to making progress on data overload. To repeat, the significance of a piece of data depends on:

- other related data,
- how the set of related data can vary with larger context,
- the goals and expectations of the observer,
- the state of the problem solving process and stance of others.

There is a widespread myth that information is something in the world that does not depend on the point of view of the observers and that it is (or is often) independent of the context in which it occurs. This is simply not the case. There are no facts of fixed significance. The available data are raw materials. A particular datum gains significance or meaning only from its relationship to the context in which it occurs or could occur including the perspective of observers. As a result, informativeness is not a property of the data field alone, but is a relationship between observers and the data field.

Take the case of a message about a thermodynamic system which states that valve X is closed. Most simply, the message signals a component status. If the operator knows (or the message also states) that valve X should be opened in the current mode of operation, then the message signals a misaligned component. Or the message could signify that with valve X closed, the capability to supply material to reservoir H via path A is compromised. Or given still additional knowledge (or data search), it could signify that with valve X closed, the process that is currently active to supply material to reservoir H is disturbed (e.g., data such as actual flow less than target flow, or no flow, or reservoir H inventory low). Furthermore, the significance of the unavailability or the disturbance in the material flow process depends on the state of other processes (e.g., is an alternative flow process available or is reservoir H inventory important in the current operating context?). Each interpretation is built around what an object affords the operator or supervisor of the thermodynamic system, including an implicit response: correctly align component, ensure capability to supply material (or take into account the consequences of the inability to do so), repair the disturbance in the material flow process (or cope with the consequences of the disturbance), or discount these messages based on other current objectives of greater importance for the context.

In this example, the significance of a datum depends on, first, a set of contextual data. Second, which pieces of data fall into this relevance set can change both with system state and with the state of the problem solving process. The latter is particularly important--what data are relevant depend on where one is in the problem solving process. Examples of how the supervisor's situation assessment or mindset affects the interpretation of an alarm include:

- If the background situation assessment is 'normal system function,' then the alarm is informative, in part, because it signals that conditions are moving into abnormal or emergency operations,
- If the background line of reasoning is 'trying to diagnose an unexpected finding,' then the alarm may be informative because it supports or contra-indicates one or more hypotheses under consideration.
- If the background line of reasoning is 'trying to diagnose an unexpected finding,' then the alarm may be informative because it functions as a cue to generate more (or to broaden the set of) candidate hypotheses that might explain the anomalous process behavior.
- If the background line of reasoning is 'executing an action plan based on a diagnosis,' then the alarm may be informative because it functions as a cue that the current working hypothesis may be wrong or incomplete since the monitored process is not responding to the interventions as would be expected based on the current working hypothesis.

Given hindsight or the position of an omniscient observer, one can specify exactly what data are needed for the ultimate solution. However, this point of view misses the cognitive task of focusing in on that relevant subset that is critical from the point of view of the person in the problem solving situation. Hindsight bias obscures the critical function in cognitive work. This is why technology-centered approaches have been unsuccessful in coping with data overload. They miss what makes this issue difficult, and they miss the opportunity to learn from how people are able extract meaning in natural fields despite being bombarded with sensory stimulation at an elemental level of analysis.

## How Are People Able To Focus In On What Is Interesting?

Since people have the ability to cope with the context sensitivity of what is informative, they become the model for how to be competent at this task, a model that we need to understand in order to make fundamental progress (and, it is important to note that people are the only extant competence model).

Mechanisms of human perception and cognition that enable people to focus on the relevant subset of the available data, even though what is interesting depends on context, include:

- processes of perceptual organization, e.g.,
  ~ pre-attentive processing that organizes the perceptual field into meaningful units and relationships,
  ~ the fact that there exist nested layers of structure in natural perceptual fields.
- processes of attentional control, e.g.,
  ~ a mix of goal-directed and stimulus-driven processing,
  ~ the center-surround structure of vision,
  ~ the relationship between focal attention and orienting perceptual functions,
- anomaly-based processing, e.g.,
  ~ contrast-based computations that pick out and focus on anomalies (departures from typicality) and that depend on relative differences (difference in a background).

Perceptual Organization

> I am standing at the window and see a house, trees, sky. And now, for theoretical purposes, I could try to count and say: there are…327 nuances of brightness [and hue]. Do I see "327"? No; I see sky, house, trees.
>
> Wertheimer, 1923 (translated from the original by N. Sarter)

The quote from Wertheimer captures a fundamental aspect of human perception and cognition that relates to data overload in virtual environments. If one counts elements in the perceptual field, there are an overwhelming number of basic elements varying in hue, saturation, and brightness across the visual field. But this avalanche of data does not overwhelm us because the processes of perception structure the scene into a few objects, events, and relationships between those objects (sky, house, trees). As one commentator on perception put it, "The process of organization reduces the stimulus data ... it groups large number of picture elements into a small number of *seen objects* and their parts" (Goldmeier, 1982, p. 5).

Meaning attaches to the end product of the grouping. The parts of the scene exist not as simply components of a larger whole, rather they act as carriers of their function within the whole. "What is perceived ... are the units and subunits, figures on a background, which result from perceptual grouping." The observer sees a field "... composed of objects, things, their form, their parts and subparts, *rather than of an enormous list of stimulus elements*" (Goldmeier, 1982, p. 5, emphasis added). The parts and elements define higher levels of structure -- objects and events in the world (Flach et al., 1995).

The ubiquitous computerization of the workplace provides the designer with the freedom to create a virtual perceptual field. The designer can (and must) manipulate the perceptual attributes of the virtual field (Wertheimer's 327+ nuances of hue, brightness, saturation and shape, motion, etc. relative to other parts of the visual scene) that would automatically specify objects and their relationships in a natural scene. This results in a need to understand how perceptual attributes and features can be used as resources whose joint effect produces an organized and coherent virtual perceptual field.

One approach (data overload results from too much "stuff") suggests that the answer to cluttered computer displays and data overload is to reduce or filter out data. Only use a few color categories. Reduce the number of pixels. Indicate less on the display. In contrast, studying how the perceptual system works in natural fields, as summarized above, leads us to a different approach. What matters in avoiding clutter and confusion is *perceptual organization*.

For example, some interface design guidelines have suggested that limits be set for optimal or maximum density, where density was defined as the number of graphical elements (pixels) versus the maximum number of locations available for graphical elements (the total pixels available in the display) -- "18% is the optimal number of CRT pixels which should be lit." However, as Wertheimer indicates, the raw density of points of luminance is not an appropriate unit of analysis from a human perception point of view (nor are they the effective stimuli). Rather, one should count in units based on what is perceived. As Hanson (1958, p. 13) put it, "the plot is not another detail in the story, nor is the tune another note."

More marks in the medium for representation, if they are used to better organize the virtual field, will reduce clutter. Tufte (1990) illustrates this approach admirably. "It is not how much empty space there is, but rather how it is used. It is not how much information [read, data] there is, but rather how effectively it is arranged" (Tufte, 1990, p. 50). Clutter results from a failure to design the elements into a coherent perceptual organization or from a failure to manipulate the elements so that the resulting perceived organization captures a meaningful organization in the referent domain. Clutter occurs when people can perceive only the perceptual attributes themselves instead of a small number of objects, their parts, and their inter-relationships in the scene. Perceptual organization (perceptual grouping and figure/ground relationships) is one critical factor in avoiding clutter and confusion. We perceive objects and events rather than elemental physical parameters of the stimuli themselves. For human perception, attributes cohere to form objects and events, and we always experience all of the perceptual attributes associated with an object.

## Control of Attention

> Everyone knows what attention is. It is the taking possession by the mind, in a clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought.

William James, 1890 I, pp. 403-404

We are able to focus, temporarily, on some objects, events, actions in the world or on some of our goals, expectations or trains of thought *while remaining sensitive to new objects or new events that may occur*.

Focus of attention is not fixed, but shifts to explore the world and to track relevant changes in the world. On flight decks, in operating rooms, and in everyday work activities, attention must flow from object to object and topic to topic. In other words, one re-orients attentional focus to a newly relevant object or event from a previous state where attention was focused on other objects or on other cognitive activities (such as diagnostic search, response planning, and communication to other agents). New stimuli are occurring constantly. Sometimes such new stimuli are distractions. But other times, any of these could serve as a signal we should interrupt ongoing lines of thought and re-orient attention. This re-orientation involves disengagement from a previous focus and movement of attention to a new focus. Interestingly, this control of attentional focus can be seen as a skillful activity that can be developed through training or supported (or undermined) by the design of artifacts and intelligent machine agents.

Thus, a basic challenge for any cognitive agent at work is where to focus attention next in a changing world. Which object, event, goal or line of thought we focus on depends on the interaction of two sets of activity. One of these is goal or knowledge directed, endogenous processes that depend on the observer's current knowledge, goals and expectations about the task at hand. The other set of processes are stimulus- or data-driven where attributes of the stimulus world (unique features, transients, new objects) elicit attentional capture or shifts of the observer's focus. These salient changes in the world help guide shifts in focus of attention or mindset to relevant new events, objects, or tasks.

The ability to notice potentially interesting events and know where to look next (where to focus attention next) in natural perceptual fields depends on the *coordination* between orienting perceptual systems (i.e., the auditory system and peripheral vision) and focal perception and attention (e.g., foveal vision). The coordination between these mechanisms allows us to achieve a "balance between the rigidity necessary to ensure that potentially important environmental events do not go unprocessed and the flexibility to adapt to changing behavioral goals and circumstances" (Folk et al. 1992, p. 1043).

The orienting perceptual systems function to pick up changes or conditions that are potentially interesting and play a critical role in supporting how we know where to look next. To intuitively grasp the power of orienting perceptual functions, try this thought experiment suggested by Woods and Watts (1997): put on goggles that block peripheral vision, allowing a view of only a few degrees of visual angle. Now think of what it would be like to function and move about in your physical environment with this handicap. Perceptual scientists have tried this experimentally through a movable aperture that limits the observer's view of a scene (e.g., Hochberg, 1986). Although these experiments were done for other purposes, the difficulty in performing various visual tasks under these conditions is indicative of the power of the perceptual orienting mechanisms.

Anomaly-Based Processing

> ... readiness to mark the unusual and to leave the usual unmarked—to concentrate attention and information processing on the offbeat.
> J. Bruner, 1990, p. 78

Another hallmark of human cognitive processing is that we tend to focus on departures from typicality (this is demonstrated at all levels of processing). We do not respond to absolute levels but rather to contrasts and change. Meaning lies in contrasts -- *some departure from a reference or expected course*.

Our attention flows to unexpected events. An event may be expected in one context and therefore go apparently unnoticed, but the same event will be focused on when it is anomalous relative to another context. An event may be an expected part or consequence of a quite abnormal situation, and therefore draw little attention. But in another context, the absence of change may be quite unexpected and capture attention because reference conditions are changing.

Our processing is tuned to contrasts -- differences in background. We process how the actual course of behavior follows or departs from reference or expected sequences of behavior given the relevant context.

## BEYOND FINESSES TO THE CONTEXT SENSITIVITY PROBLEM

> Well adapted cognitive work occurs with a facility that belies the difficulty of the demands resolved and the dilemmas balanced.
> Woods, 2001

### Typical Finesses To The Context Sensitivity Problem

At the heart of data overload is the fact that what is meaning depends on relationships among data and relationships between data and the goals and expectations of observers. Some techniques to cope with data overload may attempt to make machine reasoning more sensitive to context as an autonomous agent, while others are aimed at restructuring what is visible about virtual worlds of data to enable the basic human competence to operate as it does in natural perceptual fields. But in the end, as the field of human-computer interaction has become aware (Greenberg, 2001), no substantial progress is possible on data overload without coping in one way or another with the context sensitivity of what is informative.

Particular techniques often try to finesse the context sensitivity problem, that is, they avoid confronting the problem directly, remaining content to nibble away at it through indirect means – scale reduction, global prioritization, intelligent software agents, syntactic similarity algorithms. This leads us to an explanation for the observation that new computer based systems often exacerbate, rather than resolve, data overload. Why is this the case -- because designers try to finesse the context sensitivity problem, either avoiding or hiding how context affects what is informative.

Calling a technique a finesse points to a contrast. In one sense, a finesse is a positive pragmatic adaptation to difficulty. All of the finesses we note here are used to try to reduce data overload problems to manageable dimensions to allow experienced people to exhibit the fundamental human competence at extracting significance from data. However, a finesse is a limited adaptation because it represents a workaround rather than directly addressing the factors that make it difficult for people to extract meaning from data. In particular settings these finesses will be more or less brittle. Brittle techniques cope with some aspect of context sensitivity but break down quickly when they encounter more difficult cases.

Technology-centered approaches to data overload generally adopt strategies based on one or more of the following finesses because of inaccurate or oversimplified models of why data overload is a generic and difficult issue (for example, all of the following have been tried with some local success in coping with data overload in alarm systems; Woods, 1995a).

<u>(a) The scale reduction finesse -- reduce available data.</u>
Scaling back the available data is an attempt to reduce the amount of stuff people have to sort through to find what is significant. The belief is that if we can reduce the size of the problem, then human abilities to find the critical data as the context changes will function adequately. Often scale reduction attempts are manifested as shifting some of the available data to more "distant" secondary displays with the assumption that these items can be called up when necessary.

This approach breaks down because of the *context catch*—in some contexts some of what is removed will be relevant. Data elements that appear to be less important on average can become a critical piece of evidence in a particular situation. But recognizing their relevance, finding them and integrating them in to the assessment of the situation becomes impossible if they have been excluded or pushed into the background of a virtual data world.

This finesse also breaks down because of the narrow *keyhole catch*—proliferating more displays hidden behind the keyhole of the CRT screen creates navigation burdens (Woods and Watts, 1997). Reducing the data available on individual displays pushes data onto more displays and increasing demands for across display search and integration. This makes data available in principle, but it does not help the observer recognize or determine what would be relevant.

(b) The global, static prioritization finesse -- only show what is "important."
A related finesse is to select only the "important" subset of the available data. Often, the world of data is divided into two or three "levels of importance." Domain knowledge is used to assign individual data items to one of the two or three levels. All data items identified in the highest level of "importance" would be displayed in a more salient way to users. Data elements that fall into the second or third class of less important items would be successively less salient or more distant in the virtual world of the display system and user interface.

This approach also breaks down because of the *context catch*—how do we know what is important without taking context into account. Context sensitivity means that it is quite difficult to assign individual elements to a place along a single, static, global priority or importance dimension. Inevitably, one is forced to make comparisons between quite disparate kinds of data and to focus on some kinds of situations and downplay others. Again, data items that are not important based on some overall criteria can be critical in particular situations.

This finesse, like the first, uses inhibitory selectivity, that is, they both, in effect, throw away data. In this case, developers will object saying that users can always call up data assigned to lower levels of importance if they feel they are relevant in a particular situation. But the problem is to help people recognize or explore what might be relevant to examine without already knowing that it is relevant. To aid this process requires one to consider perceptual organization, control of attention and anomaly recognition as discussed earlier.

(c) The intelligent agent finesse -- the machine will compute what is important for you.
Another version of the *context catch* plagues this approach -- how does the machine know what is important without being able to take context into account? However, this finesse also breaks down in the face of a new catch—the *clumsy automation catch*. The observer now has another data source/team member to deal with when they can least afford any new tasks or any more data (Sarter et al., 1997).

The irony here is that developers believe that shifting the task to a computer somehow makes the cognitive challenges of focusing in on the relevant subset disappear. In fact, all finite cognitive processors face the same challenges, whether they are an individual, a machine agent, a human-machine ensemble, or a team of people. Just as machine diagnosis can err, we cannot expect machine agents to consistently and correctly identify all of the data that is relevant and significant in a particular context in order to bring it to the attention of the human practitioner. It always takes cognitive work to find the significance of data.

For example, attempts in the mid-80's to make machine diagnostic systems handle dynamic processes ran into a data overload problem (these diagnostic systems monitored the actual data stream from multiple sensors). The diagnostic agents deployed their full diagnostic reasoning power in pursuit of every change in the input data streams (see Woods, Pople, and Roth, 1990; Roth, Woods and Pople, 1992; Woods, 1994). As a result, they immediately bogged down, dramatically failing to handle the massive amounts of data now available (previously, people mediated for the computer by selecting "significant" findings for the computer to process). To get the diagnostic systems to cope with data overload required creating a front end layer of processing that extracted, out of all of the changes, which events were "significant" findings that required initiating a line of diagnostic reasoning. In this case determining what were significant events for diagnosis required determining what were unexpected changes (or an unexpected absence of a change) based on a model of what influences were thought to be acting on the underlying process.

<u>(d) the syntactic finesse -- use syntactic or statistical properties of text (e.g., word frequency counts) as cues to semantic content.</u>
This finesse is relied on heavily in keyword search systems, web search engines, and information visualization algorithms that utilize "similarity" metrics based on statistical properties of the text (e.g., frequency counts of different content words) to place documents in a visual space (e.g., Morse and Lewis, 1997; Wise, Thomas, Pennock, Lantrip, Pottier, Schur, and Crow, 1996). The primary limitation of this approach is that syntactic and statistical properties of text provide a weak correlate to semantics and domain content. There is rarely a simple one to one relationship between terms and concepts. It is frequently the case that one term can have multiple meanings (e.g., Ariane is both a rocket launcher and a proper name; ESA stands for the European Space Agency, Environmental Services Association, and the Executive Suite Association) and that multiple terms can refer to the same concept (e.g., the terms 'failed', 'exploded', and 'was destroyed' can be used interchangeably).

The problem is compounded by the fact that the 'relevance' metrics employed (e.g., the weighting schemes used by web search engines) are often opaque to the user. This is the *lack of observability* catch. The user sees the list of documents retrieved based on the query and the relevance weighting generated by the search engine. However, in many cases how the relevance weighting was generated is unclear, and the resulting document ordering does not accord well with how the user would have prioritized the documents (i.e., documents that come up early with a high weighting can be less relevant than documents that come up later). This forces the user to resort to attempting to browse through the entire list. Since the generated list is often prohibitively long, it can leave the user unsure about whether important documents might be missed. Users will often prefer to browse documents ordered by metrics that do not attempt or claim to capture "relevance," such as date or source, rather than by syntactic relevance weighting because the organizing principle is observable and they know how to interpret values along those dimensions.

Attempts to place documents in a visual space based on syntactic properties are also subject to the *over-interpretation* catch. The spatial cues and relationships that are visible to the observer will be interpreted as meaningful even if they are incidental and not intended to be information bearing by the designer (or algorithm). For example, visualizations that attempt to represent multi-dimensional spaces (4 or more dimensions) on a two dimensional display can create ambiguities with respect to the position of a document relative to each of the dimensions. Users may assume that two documents that are located close to each other on the display reflect a similar degree of relationship to each of the dimensions represented in the space, when in fact they are not in the same position in the multi-dimensional space – even though it looks that way on the display.

## CONSTRAINTS ON EFFECTIVE SOLUTIONS TO DATA OVERLOAD

We can now compactly summarize our diagnosis:

- Why is data overload so difficult to address? Context sensitivity – meaning lies, not in data, but in relationships of data to interests and expectations..
- Why have new waves of technology exacerbated, rather than resolved, data overload? When they ignore or finesse context sensitivity.
- How are people able to cope with data overload? People are able to shift focus of attention fluently as circumstances change and re-orient to potentially interesting new stimuli.

This diagnosis, based on people's natural ability as a model of competence, generates a number of constraints on effective solutions to data overload.

*1. **Organization** precedes selectivity.*

Solving data overload begins with creating a structured field of view on which attention (selectivity) can operate, focusing on potentially interesting areas depending on context. Designers will need to define the groups/objects/events and relationships attention can select.

The default in computer systems has been to organize around elemental data units or on the units of data appropriate for computer collection, transmission, and manipulation (Flach et al., 1995). These are either too elemental, as if we saw the world in 'Werthheimer's 327+' variations in hue, saturation, and brightness, or too removed from the meaningful objects, events and relationships for the user's field of practice.

This finding means that effective systems for coping with data overload

- will have elaborate indexing schemes that map onto models of the structure of the content being explored.
- will need to provide multiple perspectives to users and allow them to shift perspectives fluently.

*2. Positive selectivity* ***enhances*** *a potion of the structured field.*
All approaches to data overload involve some sense of selectivity. However, there are different forms of selectivity: facilitation or inhibition of processing. In the former, selectivity *facilitates* or enhances processing of a portion of the whole. In this form of selectivity, we use positive metaphors such as a spotlight of attention or a peaked distribution of resources across the field.

In the latter, selectivity *inhibits* processing of non-selected areas, for example stimuli in the selected portion can pass through and go on for further processing, whereas stimuli in the non-selected portion do not go on for processing. In this form of selectivity, we use negative metaphors such as a filter or a gatekeeper.

Commonly, the response to data overload is filtering, that is, negative selectivity. But the surprise is that this is not how people solve data overload because it undermines or eliminates the ability to switch focus. The critical criterion for processes of selection, parallel to human competence, is that observers *need to remain sensitive to non-selected parts in order to shift focus fluently as circumstances change or to recover from missteps*. This means that human attentional focus is not a limited channel to be overcome with prostheses, but rather a necessary part of being sensitive to context which should be emulated.

This finding means that effective systems for coping with data overload will use positive forms of selectivity and develop techniques that support shifting focus over the field of data.

*3. All techniques to cope with data overload must deal with* ***context sensitivity****.*
Data are informative based on *relationships* to other data, relationships to larger frames of reference, and relationships to the interests and expectations of the observer. Making data meaningful always requires cognitive work to put the datum of interest into the context of related data and issues.

This finding means that solutions to data overload will help practitioners put data into context. Presenting data in context shifts part of the burden to the external display rather than requiring the observer to carry out all of this cognitive work "in the head." Many techniques could support this criterion (Woods, 1995b). One, when we display a given datum, we can show it in the context of related values. Second, rather than organizing displays around pieces of data, we can organize data around meaningful issues and questions--model based displays. These are models of how data relationships map onto meaningful objects, events, and processes in the referent field of activity (Flach et al., 1995).

Three, we can use the power of the computer to help extract events from the flow of elemental data (Christoffersen, Woods and Blike, 2001). Events are temporally extended behaviors of the device or process involving some type of change in an object or set of objects. Fourth, the computer could also help observers recognize anomalies and contrasts by showing how the data departs from or conforms to the contrasting case (a departure from what is expected, from what is the plan or doctrine, from what has been typical). Since there are usually many possible contrasting cases, each defines a kind of perspective around which one views the available elemental data.

There is a prerequisite for the designer to be able to put data into context: they need to know what relationships, events, and contrasts are informative over what contexts in the field of practice.

*4. **Observability** is more than mere data availability.*

> The greatest value of a picture is when it forces us to notice what we never expected to see.
> Tukey, 1977, p. vi

There are significant differences between the available data and the meaning or information that a person extracts from that data. *Observability* is the technical term that refers to the cognitive work needed to extract meaning from available data (Rasmussen, 1985). This term captures the relationship among data, observer and context of observation that is fundamental to effective feedback.

Observability is distinct from data availability, which refers to the mere presence of data in some form in some location. For human perception, "it is not sufficient to have something in front of your eyes to see it" (O'Regan, 1992, p.475). Recent studies on visual perception emphasize the role of attention in the perception of suprathreshold stimuli (Resnick, O'Regan, and Clark, 1997; Simons and Levin, 1997)-- "there seems to be no conscious perception without attention." (Mack and Rock, 1998, p.ix).

One example of displays with very low observability occurs on the current generation of flight decks. The flight mode annunciations are a primary indication of how automated systems are configured to fly the aircraft. These crude indications of automation activities contribute to automation surprises where the automation flies the aircraft in a way that the pilots did not anticipate. As one pilot put it, "changes can always sneak in unless you stare at it" (see Woods and Sarter, 2000, for more on this example, and Sarter, 2000 for tests of different approaches to enhance observability in this context).

Observability refers to processes involved in extracting useful information. It results from the interplay between a human user knowing when to look for what information at what point in time and a system that structures data to support attentional guidance (see Rasmussen, 1985; Sarter, Woods and Billings, 1997). *The critical test of observability is when the display suite helps practitioners notice more than what they were specifically looking for or expecting*. If a display only shows us what we expect to see or ask for, then it is merely making data available.

*5. To cope with data overload, ultimately, requires the design of **conceptual spaces**.*
One builds a conceptual space *by depicting relationships in a frame of reference* (Woods, 1995b; Rasmussen et al., 1994). The search to solve data overload begins with the search for frames of reference that capture meaningful relationships for that field of practice. A frame of reference is a fundamental property of a space and what makes a space or map special from the point of view of representation. With a frame of reference comes the potential for concepts of neighborhood, near/far, sense of place, and a frame for structuring relations between entities. A frame of reference is a prerequisite for depicting relations rather than simply making data available.

Almost always there are multiple frames of reference that apply. Each frame of reference is like one perspective from which one views or extracts meaning from data. Part of designing a conceptual space is discovering the multiple potentially relevant frames of references and finding ways to integrate and couple these multiple frames.

## TOWARD CONTEXT-SENSITIVE APPROACHES

The preceding discussion leads to the conclusion that, because of the context-sensitivity problem, we must direct our efforts towards techniques which do not rely on knowing in advance what subset of data is relevant (Greenberg, 2001). We have also argued that methods which rely centrally on machine processing are vulnerable to brittleness. This has led us to cast the problem as one of helping people to recognize or explore the portions of the data field that might be relevant so that they can shift focus attention to those areas as the current situation evolves.

We believe a context bound approach to data overload dictates two parallel strategies to move toward more effective solutions. The first is to use models of the domain semantics and content as the foundation for visualizations which provide a structured view of the potentially meaningful *relationships* across data for observers, particularly, events and contrasts (Flach et al, 1995). The intent of *model-based conceptual spaces* is to take advantage of the context-sensitive properties of human cognition by giving observers the perceptual leverage needed to focus in on relevant sub-portions of the data space.

This tactic is similar in philosophy to other methods in the literature which use models of domain semantics as a way to structure displays of data (e.g., Vicente and Rasmussen, 1992). Taking advantage of the context-sensitive nature of human cognition presumes a structured data field on which our attentional processes operate. The idea therefore is to build a conceptual space for organizing the data based on a model of the fundamental relationships, objects, and events in the domain. In order to support skillful shifting of attention, these visualizations will have to include mechanisms allowing observers to perceive changes or potentially interesting conditions which are not necessarily in direct view, and to re-orient their attention to the new data. They must also emphasize anomalies and contrasts by showing how data departs from or conforms to expectations.

Different kinds of models become the basis for different conceptual spaces that depict different sets of relationships potentially relevant to the field of activity. These become sets of perspectives practitioners can shift across analogous to shifting point of view in exploring a physical space. Patterson, Woods, and Tinapple (2001) provide an example of coordinating multiple conceptual spaces in the context of analysis tasks to illustrate the general strategy.

Because of the central importance of models in this method, it is important to develop a framework of models which can be used to understand different settings and to serve as the basis for choosing and instantiating appropriate classes of visualizations. Levels of the framework will vary in terms of how closely bound they are to a particular domain or scenario. Although there will be models that will generalize across all domains such as models of function or of events, a context-bound approach implies that some models will be instantiated that are more specific (Patterson, Roth and Woods, in press).

A conceptual space to structure the field of data and allow people's natural competencies at context sensitive processing to emerge, then, provides the base on which one can build cooperative roles, either human-machine or human-human. The model-based conceptual spaces function as the kind of common ground necessary to all forms of collaborative work (Clark and Brennan, 1991; Olson and Olson, 2000).

In this context, active machine intelligence can play supplemental, circumscribed, cooperative roles to aid human observers in organizing and re-organizing data within and across multiple model-based perspectives. These roles would fulfill the requirement for using machine intelligence in ways that relax the need for these algorithms to be correct almost always.

Patterson, Woods, and Tinapple (2001) illustrate this approach around the concept of 'source quality' in inferential analysis. Rather than having the machine intelligence sort by a summary, integrative judgment such as source quality, the dimensions and criteria used to reach such judgments could be used to organize the display of the data, i.e., they form a model of what a high level concept like 'source quality' means. A conceptual space makes that model visible so that one sees source documents structured around these differing dimensions. Practitioners then can use their knowledge of the situation to trade off tensions, gaps, and uncertainties in the visible model, sensitive to the context that eludes literal minded machines. In addition, machine intelligence can play an active critiquing role against this base by suggesting ways to broaden the exploration of the data space to see other patterns and by raising alternative hypotheses to explain the visible patterns (e.g., Guerlain et al., 1999).

**Criteria for Innovation**

The promise of new technology is more than making data available. New technology does provide the power to develop external support for the cognitive activities involved in extracting the significance from data. The question is how to use that power. Ironically, this power can be used (and has been) to exacerbate data overload as well as to support people's ability to interpret large fields of data.

The diagnosis presented here points to criteria and constraints that need to drive an innovation process. Rather than rely on finesses, innovation needs to tackle the central role of context sensitivity. The basic human competence for finding what is informative in natural perceptual fields despite context sensitivity can serve us well as a guide in the innovation process. Drawing on human competence teaches us several surprising lessons: organization precedes selectivity; positive forms of selectivity preserve the ability to re-focus attention; capturing relationships as the fundamental unit of display allow escape from the syndromes of elemental data-- Werthheimer's 327+ nuances and James' baby; coordinating orienting and focal perceptual functions allows us to see more than just what we were specifically looking for. Vigorous pursuit of model-based conceptual spaces can lead us to discover ways to make virtual perceptual spaces support the forms of expertise people exhibit in natural perceptual fields.

Ultimately, solving data overload problems requires both new technology and an understanding of how systems of people supported by various artifacts extract meaning from data. Our design problem is less -- can we build a visualization or an autonomous machine, and more -- what would be useful to visualize and how to make automated and intelligent systems team players. A little more technology, by itself, is not enough to solve generic and difficult problems like data overload, problems that exist at the intersections of cognition, collaboration, and technology.

**ACKNOWLEDGMENTS**

This work, as the first step in a project to innovate new concepts to escape form data overload, was made possible through the efforts of Gil Kuperman. He recognized the need to step back from the normal project and production constraints to re-examine why data overload has proven such a persistent and difficult problem. His foresight to make room for this effort was instrumental in opening up new avenues to study how people cope with data overload and to innovate new design concepts.

**REFERENCES**

Bower G. and Morrow, D. G. (1990). Mental models in narrative comprehension. *Science*, *24*, 44-48.

Brann, D.B., Thurman, D.A., Mitchell, C.M. (1996). Human interaction with lights-out automation: a field study. In *Human Interaction with Complex Systems '96*, Dayton, OH.

Bruner, J. (1990). *Acts of Meaning*. Cambridge MA: Harvard University Press.

Christoffersen, K., Woods, D. D. and Blike, G. T. (2001). Extracting Event Patterns From Telemetry Data. Proceedings of the Human Factors and Ergonomics Society 45th annual meeting. 8-12 October, Minneapolis, MN.

Clark, H. H., and Brennan, S. E. (1991). Grounding In Communication. In L. Resnick, J. M. Levine, and S. D. Teasley (Eds.), *Perspectives on Socially Shared Cognition* (pp. 127-149). Washington, D. C.: American Psychological Association.

Conan Doyle, A. (1986). *The Reigate Squire*. Hardwick (Originally published in both *Strand* and *Harper's* in June 1893).

Fischer, G. and Reeves, B. (1992). Beyond intelligent interfaces: exploring, analyzing, and creating success models of cooperative problem solving. *Journal of Applied Intelligence, 1,* 311-332.

Flach, J., Hancock, P., Caird, J. and Vicente, K. (Eds.) (1995). *An Ecological Approach To Human Machine Systems I: A Global Perspective*. Hillsdale, NJ: Lawrence Erlbaum.

Folk, C. L., Remington, R. W. and Johnston, J. J. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1030-1044.

Goldmeier, E. (1982). *The Memory Trace: Its Formation And Its Fate*. Hillsdale, N.J.: L. Erlbaum Associates.

Glymour, C. (1987). *Discovering causal structure: artificial intelligence, philosophy of science, and statistical modeling.* Orlando: Academic Press.

Greenberg, S. (2001). Context as a Dynamic Construct. *Human-Computer Interaction,* 16, in press.

Guerlain, S., Smith, P.J., Obradovich, J. Heintz, Rudmann, S., Strohm, P. Smith, J.W., Svirbely, J., and Sachs, L. (1999). Interactive Critiquing as a Form of Decision Support: An Empirical Evaluation. *Human Factors 41(1),* 72-89.

Gruen, D., Sidner, C., Boettner, C., and Rich, C. (1999). A collaborative assistant for email. In *CHI 99 ACM Conference on Human Factors in Computing Systems*, New York: ACM Press. Pittsburgh, PA. 196-197.

Hanson, N. R. (1958). *Patterns of Discovery*. Cambridge England, Cambridge University Press.

Hardwick, M. (1986). *The Complete Guide To Sherlock Holmes*, pp. 86-87. New York: St. Martin's Press.

Hochberg, J. (1986). Representation of motion and space in video and cinematic displays. In K. Boff, L. Kaufman, and J. Thomas (Eds.*), Handbook of Perception and Human Performance* (pp. 22-1 to 22-64). New York: John Wiley and Sons.

Hollnagel, E., Bye, A. & Hoffmann, M. (2000). Coping with complexity: Strategies for information input overload. Proceedings of the Conference on Cognitive Systems Engineering in Process Control (CSEPC 2000), November 22-24, Taejon, South Korea.

James, W. (1981). *The Principles of Psychology*. New York, H. Holt and Company [original 1918].

Joyce, J. P. and Lapinski, G. W. (1983). A history and overview of the safety parameter display system concept. *IEEE Transactions on Nuclear Science*, NS-30.

Kraus, J. (1979). We Wait and Wonder. *Cosmic Search* Vol. 1, No. 3 (http://www.bigear.org/vol1no3/wonder.htm)

Letsche, T.A. and Berry, M.W. (1997). Large-scale information retrieval with latent semantic indexing. *Information Sciences*, 100, 105-137.

Mack, A. and Rock, I. (1998). *Inattentional Blindness*. Cambridge, MA: MIT Press.

Maes, P. (1998) Agents that reduce work and information overload. In M.T. Maybury, W. Wahlster. *Readings in Intelligent User Interfaces*. Morgan Kaufmann Publishers.

Maes, P. and Schneiderman, B. (1997). Direction manipulation vs. interface agents: a debate. Interactions, Vol. IV Number 6, ACM Press.

Marx, M., and Schmandt, C. (1996) CLUES: Dynamic Personalized Message Filtering. In *CSCW '96 Proceedings*. Boston, MA. 113-121.

Miller, J. G. (1960). Information input overload and psychopathology. *American Journal of Psychiatry*, 116:695--704.

Morse, E and Lewis, M. (1997). Why information retrieval visualizations sometimes fail. In *Proceedings of the 1997 IEEE International Conference on Systems, Man, and Cybernetics*, Oct. 12-15, 1997, pg. 1680 - 1685.

Murray, C. and Cox, C. B. (1989). *Apollo: The Race to the Moon*. New York: Simon and Schuster.

Oakes, M.P., and Taylor, M.J. (1998). Automated assistance in the formulation of search statements for bibliographic databases. *Information Processing and Management*, 34 (6), pp. 645-668.

Olson, G. M. and Olson J. S. (2000). Distance Matters. *Human-Computer Interaction*. 15, 139—178.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology, 46,* 461-488.

Patterson, E. S., Roth, E. M. and Woods, D.D. (in press). Predicting Vulnerabilities in Computer-Supported Inferential Analysis under Data Overload. *Cognition, Technology and Work*.

Patterson, E.S., Woods, D.D., Tinapple, D., Roth, E.M. (2001). Using cognitive task analysis (CTA) to seed design concepts for intelligence analysts under data overload. Proceedings of the Human Factors and Ergonomics Society 45th annual meeting. 8-12 October, Minneapolis, MN.

Rabbitt, P. (1984). The control of attention in visual search. In R. Parasuraman and D. R. Davies (Eds.), *Varieties of Attention.* New York: Academic Press.

Rasmussen, J. (1985). Trends in human reliability analysis. *Ergonomics*, *28*(8), 1185-1196.

Rasmussen, J., Pejtersen, A. M. and Goldstein, L. P. (1994). *Cognitive Systems Engineering*. New York: John Wiley and Sons

Reed, E. S. (1988). *James J. Gibson and the Psychology of Perception*. New Haven, CT: Yale University Press.

Resnick, R. A., O'Regan, J. K. and Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8*, 368-373.

Roth, E. M., Woods D. D., and Pople, H. E. Jr. (1992). Cognitive simulation as a tool for cognitive task analysis. *Ergonomics, 35,* 1163–1198.

Sarter, N.B. (2000). The Need for Multi-sensory Feedback in Support of Effective Attention Allocation in Highly Dynamic Event-Driven Environments: The Case of Cockpit Automation. *International Journal of Aviation Psychology*, 10(3), 231-245.

Sarter, N., Woods, D. D. and Billings, C. (1997). Automation surprises. In G. Salvendy, (Ed.) *Handbook of Human Factors/Ergonomics* (2nd ed.). New York: John Wiley and Sons.

Simon, H. A. (1976). *Administrative Behavior* (3rd ed.). New York: Free Press.

Simon, H. A. (1981). *Sciences of the Artificial* (2nd ed.). Cambridge, MA: MIT Press.

Simons, D. J. and Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences, 1,* 261-267.

Smith, P.J., McCoy, E., and Layton, C. (1997). Brittleness in the design of cooperative problem-solving systems: the effects on user performance. *IEEE Transactions on Systems, Man, and Cybernetics, 27(3),* 360-370.

Tufte, E. R. (1990). *Envisioning Information*. Chesire, CT: Graphics Press.

Tukey, J.W. (1977). *Exploring Data Analysis*. Reading Massachusetts: Addison-Wesley.

Vicente, K. J. and Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *IEEE Transactions on Systems, Man, and Cybernetics*, *22(*4), 589-606.

Wertheimer, M. (1950, originally published in 1923). Laws of Organization in Perceptual Forms. In W. D. Ellis (Ed.), *A Source Book of Gestalt Psychology*. New York: Humanities Press.

Wise, J. A., Thomas, J. J., Pennock, K., Lantrip, D., Pottier, M., Schur, A., Crow, V. (1996) Visualizing the non-visual: Spatial analysis and interaction with information from text documents. *Proceedings of Info Viz 96*.

Wolfe, J. M. (1992). The parallel guidance of visual attention. *Current Directions in Psychological Science, 1*, 124-128.

Woods, D. D. (1984). Visual momentum: A concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies, 21*, 229-244.

Woods, D. D. (1991). The cognitive engineering of problem representations. In G. R. S. Weir and J. L. Alty, (Eds.) *Human-Computer Interaction and Complex Systems*. London: Academic Press.

Woods, D. D. (1993). The price of flexibility in intelligent interfaces. *Knowledge-Based Systems*, *6*, 1-8.

Woods, D. D. (1994). Cognitive demands and activities in dynamic fault management: abduction and disturbance management. In N. Stanton (ed.) *Human Factors of Alarm Design*. London: Taylor and Francis.

Woods, D. D. (1995a). The alarm problem and directed attention in dynamic fault management. *Ergonomics*, *38*(11), 2371-2393.

Woods, D. D. (1995b). Towards a theoretical base for representation design in the computer medium: ecological perception and aiding human cognition. In J. Flach, P. Hancock, J. Caird, and K. Vicente, (Eds.) *An Ecological Approach To Human Machine Systems I: A Global Perspective*, Hillsdale, NJ: Lawrence Erlbaum.

Woods, D. D. (2002). Laws that Govern Cognitive Work. Plenary Address, Annual Meeting of the Cognitive Science Society, August 2002.

Woods, D. D., Johannesen, L., Cook, R. I. and Sarter, N. B. (1994). *Behind Human Error: Cognitive Systems, Computers and Hindsight*. Wright-Patterson AFB, OH: Crew Systems Ergonomic Information and Analysis Center (CSERIAC).

Woods, D. D., Pople, H.E. Jr. and Roth, E. M. (1990). *The Cognitive Environment Simulation as a Tool for Modeling Human Performance and Reliability,Volumes I and II* (Technical Report NUREG-CR-5213). Washington D.C.: U.S. Nuclear Regulatory Commission.

Woods, D. D. and Patterson, E. (2000). How Unexpected Events Produce an Escalation of Cognitive and Coordinative Demands. In *Stress Workload and Fatigue*. P. A. Hancock and P. Desmond (eds.) Lawrence Erlbaum, Hillsdale NJ.

Woods, D. D. and Sarter, N. (2000). Learning from automation surprises and going sour accidents. In N. Sarter and R. Amalberti (Eds.) *Cognitive Engineering in the Aviation Domain*. Hillsdale NJ: Lawrence Erlbaum

Woods, D. D. and Watts, J. C. (1997). How not to have to navigate through too many displays. In Helander, M. G., Landauer, T. K. and Prabhu, P. (Eds.) *Handbook of Human-Computer Interaction, 2nd edition*. Amsterdam: Elsevier Science.

Zhang, J. and Norman, D.A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, *18*, 87–122.