# Machine Learning, Reasoning, and Intelligence in Daily Life: Directions and Challenges

**Eric Horvitz**
Microsoft Research
Redmond, Washington USA 98052
horvitz@microsoft.com

## Abstract

Technical developments and trends are providing a fertile substrate for creating and integrating machine learning and reasoning into multiple applications and services. I will review several illustrative research efforts on our team, and focus on challenges, opportunities, and directions with the streaming of machine intelligence into daily life.

## 1 Reflections on Trends and Directions

Over the last decade, technical and infrastructural developments have come together to create a nurturing environment for developing and fielding applications of machine learning and reasoning—and for harnessing automated intelligence to provide value to people in the course of their daily lives. These developments include (1) technical advancements in machine learning and reasoning, (2) the growth in CPU and memory capabilities within commonly available devices and platforms, (3) the connectivity, content, and services provided by the evolving Web, and (4) the increasing availability of data resources, including corpora of behavioral data collected via inexpensive sensors, and through ongoing interaction with software and services.

### 1.1 Panoply of Applications and Services

Opportunities for integrating applications of machine intelligence into the daily lives of people are growing with the increasing popularity of computing systems, the widening diversity of web services, the growing popularity of portable devices that contain general-purpose operating systems, and ongoing innovations in human-computer interaction—including the increasing prowess of speech recognition, handwriting, and sketch-understanding interfaces.

Various examples of the integration of automated learning and reasoning into daily life have been appearing as implicit and explicit extensions to traditional systems and services, and also in prototypes and systems that provide qualitatively new kinds of experiences. I will review several projects and efforts undertaken by our team that highlight directions and approaches to introducing potentially valuable machine learning and reasoning into the daily lives of people.

An example of an implicit integration of ambient learning and reasoning is the effort by our team to create a probabilistic action prediction and prefetching subsystem that is embedded deeply in the kernel of Microsoft's Windows Vista operating system. The predictive component, operating within a component in the Vista operating system called Superfetch, learns by watching sequences of application launches over time to predict a computer user's application launches. These predictions, coupled with a utility model that captures preferences about the cost of waiting, are used in an ongoing optimization to prefetch unlaunched applications into memory ahead of their manual launching. The implicit service seeks to minimize the average wait for applications to be ready to use after launching actions.

Moving from the depths of operating system kernels to the infrastructure of a city, there is great opportunity for collecting data and learning predictive models from constellations of sensors embedded throughout a large-scale region. The JamBayes traffic forecasting service [Horvitz *et al.*, 2005] serves as an example of an explicit extension of ambient intelligence to familiar views of digital maps that display the flow of traffic in urban areas. The JamBayes client, operating on desktop systems and portable devices, accesses the JamBayes predictive traffic service that employs machine learning and reasoning about the context-sensitive flow of traffic. The system overlays predictions about *future* traffic conditions on a digital traffic flow map. The system combines multiple variables that consider key contextual evidence as well as the dynamics of flow across a greater urban area to predict when free-flowing traffic will likely become jammed and how long it will be until current jams melt away. A descendant of JamBayes developed by our team named ClearFlow provides users with context-sensitive routing based on reasoning about the current and future flows on road segments within a regional traffic system.

Beyond serving as an overlay of predictive services on traffic flow maps, JamBayes also provides a qualitatively new functionality—*surprise detection* and *surprise forecasting*. We have been pursuing research on surprise forecasting, aimed at developing and fielding sensing and reasoning systems that have the ability to detect and to alert users when *current or future events will likely surprise them*. Such reasoning promises to provide significant value to people in the course of daily life as it explicitly considers the misalignment between computational forecasts about the world and the inferred expectations of people with regards to important outcomes. JamBayes' surprise modeling considers a user's context-sensitive expectations about current and future traffic, and infers when situations that it is inferring would likely be interpreted as anomalous or "surprising" from the user's perspective. Users can configure the system to generate alerts when particular kinds of surprises occur, based on context. In operation, the surprise modeling and analysis is used to let people know when current or projected flows (both high and low) on routes that they have expressed interest in would likely surprise them.

Let us now move to systems that learn and reason about a computer user's activities and that can provide qualitatively new experiences for users. In the Web Montage prototype [Anderson and Horvitz, 2002], data collected in the background about a user's desktop activity is analyzed via machine learning, and predictive models are constructed that are used to generate *personalized adaptive portals*, pages that are composed dynamically by coalescing automated clippings from multiple web sources. The system builds models of the cost of navigation and the value of content to triage information based on time of day and ongoing activities, including for example, the current and recent patterns of topics at focus of attention on the desktop. Content drawn in a selective manner from multiple sources is arranged and laid out via an optimization of expected value.

Another example of overlaying intelligence to provide new kinds of experiences is the MemoryLens effort [Horvitz *et al.*, 2004], which has been exploring the use of machine learning and reasoning to model aspects of human memory. The LifeBrowser prototype developed within the MemoryLens effort employs inferences about the memorability of events and items to provide new kinds of browsing and searching experiences such as variable density timelines. It also uses the inferences about landmarks to extend familiar search functionality by integrating a backbone of memorable personal events and activities into the display of time-sorted search results.

Other examples of new functionalities and services leverage components that reason about peoples' focus of attention and interruptability [Horvitz *et al.*, 1999; Horvitz *et al.* 2003b; Fogarty *et al.* 2005; Iqbal and Bailey, 2006], as well as about people's current and future presence and availabilities [Horvitz *et al.*, 2002]. We have explored the use of these components in systems that can reason about both the urgency of incoming communications and the status of a user's workload and focus of attention. Such systems have the ability to weigh the cost of deferring communications with the cost of interrupting the user. Examples of prototypes constructed in this realm include the Priorities system [Horvitz *et al.*, 1999], the larger cross-platform alerting system named Notification Platform [Horvitz *et al.* 2003], and several of our Bestcom efforts. Research in this realm highlights the potential for machine learning and reasoning to play a significant role with enhancing communications.

The Notification Platform continues to build and triage information from multiple sources within a unified inbox, and reasons about the best devices and modalities to employ. The related Bestcom effort (for *best-means communications*) centers on assisting people to establish communications with one another. Within Bestcom research, we have investigated the promise of systems with the ability to understand communication preferences, the goals of a communication, current and future availabilities, and context. Prototypes explore how agents, working on behalf of a contactor and contactee, can negotiate about the best time and type of communication to undertake, and then execute connections. In the general Bestcom methodology, a contactor requests a communication with a contactee (the contactor "Bestcoms" the contactee), and an optimization is executed based on identity, context, goals, connectivity, and devices available now and in the future. Actions may range from a real-time voice call to a rescheduling of a voice or video-conference for a later time. Various prototypes have been created, including systems that perform smart routing of calls based on context, executing on a portable device (*e.g.*, the Bayesphone [Horvitz, *et al.*, 2005]), and systems employing a larger infrastructure, such as the Bestcom-ET system fielded internally at Microsoft [Horvitz *et al.*, 2003c].

Machine learning and reasoning can provide value with the overall coordination of people in the course of daily life. The Coordinate system [Horvitz *et al.*, 2002] was initially developed to support the Notification Platform and Bestcom efforts, but the work also led to standalone presence and availability forecasting services that provide people with new kinds of awareness. Coordinate continues to collect data about the presence and activities of people at different locations and devices, and employs machine learning and reasoning to perform availability forecasting. The system can pass its inferences to computational agents, or provide to people (who have been granted access privileges), such predictions as the time until a user will return to their office, read email, regain access to a networked computing system, or finish a conversation that is currently in progress.

### 1.2 Rise of Intention and Preference Machines

Moving from particular examples to trends, we are seeing the use of machine learning, inference, and decision making to drive the creation of *preference machines* and *intention machines* in multiple domains. Preference machines include the set of systems referred to as recommender systems, employing collaborative filtering to predict the preferences of users about different sets of items, content, and experiences based on partial information about activities, demographics, and preferences. Such systems typically leverage predictive models constructed from the activities or assessments of a large set of users. The world has come to know of such systems as recommendations in online commerce situations.

Intention machines are services that employ models that predict peoples' activities and goals. Such work includes the predictive models used in the Microsoft operating system described earlier, and numerous other projects. As an example of intention machines in common use, the lives of many millions of people are touched daily via web search engines that reason about the goals of users given sparse queries. Logs of queries and page accesses serve as a rich case library for building predictive models. Microsoft has been investigating the use of machine learning based on logs of user search activity to continue to enhance the functions used to rank results associated with queries. Recent work on constructing predictive models from logs of user data is highlighted in [Burges *et al.*, 2005] and [Downey *et al.*, 2007]. Other models are used to identify overall satisfaction with results. Learning and reasoning is also being used to optimize the presentation of advertisements. Such research is undoubtedly going on at other companies providing Web search, and services associated with targeted advertising.

Intention and preference machines will see even more exotic uses, including their use in geocentric services. As an example, our team has been building intention and preference machines with data we collected over several years from volunteers who participated in the Microsoft Multi-User Location Survey (MSMLS). We have been exploring the uses of the data in learning and reasoning systems, including the construction of a system that can predict and then harness drivers' likely destinations, given initial driving trajectories [Krumm and Horvitz, 2006]. Beyond geocentric intention machines, we have been exploring the feasibility of building geocentric preference machines, that perform *geocentric collaborative filtering*: Given sets of sensed destinations of multiple people and the sensed destinations of a particular driver, what places, unvisited previously by that driver, might be of interest, and how and when might the driver be best informed (*e.g.*, by hearing a paid advertisement when he or she is approaching such destinations).

Intention machines and preference machines are becoming important and increasingly common assets in business. Competitive pressures will lead to an ongoing polishing of these models and increasingly elegant, desirable interactions with people.

## 2 Challenges and Opportunities

Bringing the fruits of machine intelligence into daily life faces an array of interesting challenges—and with the challenges come opportunities for innovation.

### 2.1 Learning and Supervision

Designs for introducing intelligent reasoning into the world often depend critically on acquiring a case library of rich data that can be used to build predictive systems. The construction of case libraries for machine learning often requires the labeling of hidden states, such as the ground truth of intentions or preferences. In some situations, users may have to engage in explicit tagging activities or *supervision*. In many cases, it is possible to tag intentions and preferences with *in-stream supervision*, a phrase we use to refer to assigning labels in an implicit manner, in the course of activity. In-stream supervision includes situations where the target states of interest are tagged according to definitions of logged events. As an example, the core predictive models of Coordinate employ in-stream supervision, where states and transitions that capture user presence and availability are logged automatically. Web Montage and Vista Superfetch harness in-stream supervision in a similar manner.

As another example, in-stream supervision employed in the Lookout calendaring and scheduling agent [Horvitz, 1999], is used to build two predictive models without cost to the user. Lookout computes the likelihood that someone reviewing email might like to perform a scheduling task based on the content of the email message at focus. The system collects and labels cases for building a model of users' intentions in the course of a user's normal activity. To collect cases, Lookout runs in the background and notes when users examine an email message and then turn, within a time horizon, to a calendar view or scheduling task. Positive and negative examples of messages and the details of message headers and bodies are stored along with the observed action in a case library. Lookout also uses in-stream supervision to learn about the ideal timing of actions. The system watches behind the scenes and records the amount of time that people focus on email messages before moving onto calendaring tasks or onto other messages. Lookout builds a case library of messages and dwell times collected through such in-stream supervision and constructs a predictive model that provides real-time recommendations about how long the system should wait before engaging the user, given properties of the message (such as the message length) at the user's focus of attention. Such a predictive model, learned without human intervention, provides Lookout with an awareness of attentional patterns of users, enabling Lookout to courteously withhold potentially distracting engagements while a user reviews a message.

A similar approach to supervision is used in the Priorities system for prioritizing email by urgency. Priorities learns to infer the expected cost of delayed review for each incoming email message. Predictive models are constructed based on explicit supervised learning or on in-stream supervision, considering a user's activities. For an example of the latter, when running in in-stream supervision mode, messages that are deleted without being read are assigned a lower urgency than messages read soon after they arrive.

In-stream supervision methods do not have to be fully automatic and operate as complete sleuths. Priorities research explored a middle ground of allowing users to become more involved with in-stream supervision. In versions of Priorities, users could inspect and modify in-stream supervision policies. Such awareness and potential modification allows the in-stream supervision to become a grounded collaboration between the machine and user. The system also allows users to inspect the case libraries before invoking the modification or regeneration of predictive models. By collaborating in such a mixed-initiative manner *about* the process of the tagging process, in-stream supervision can be made more accurate.

A number of systems have employed more costly probes for hidden states. For example, the BusyBody system learns and then uses personalized models that predict the cost of interruption of a computer user based on the user's activity and context [Horvitz *et al*., 2004b]. During a training phase, BusyBody makes intermittent requests for the current cost of interruption. We have investigated minimizing user training effort by automatically tagging a user's cost of interruption via a proxy such as the delays in responding to notifications. Developing a grounded collaboration on learning via using a sharing of policies is promising.

Beyond in-stream supervision, unsupervised and semi-supervised learning show promise for reducing the cost of training systems. In another approach, we have been exploring the harnessing of active learning in guiding the allocation of supervision efforts, including an approach we refer to as *selective supervision*—the application of decision-theoretic methods to triaging cases for explicit tagging [Kapoor *et al.,* 2007]. This research includes efforts on *lifelong learning* focused on developing systems that reason in an ongoing manner about the costs of additional probing of users for input versus the long-term benefits of enhanced performance of a system in an environment over time.

On a related challenge, people may wish to use a system right away, before training a system. In one approach, pre-trained *generic* models can be made available and users can select and use the most appropriate model immediately. Such generic models may often not be able to provide the accuracy delivered by a personalized model. However, they can deliver value immediately. The case libraries of such models can be extended or washed out over time with new user-specific data. In another approach, a model mixture approach can be used, where the output of the pre-trained model is combined with the output of a personal model that is growing more sophisticated over time with addition of new labeled cases—and the fusion of the predictions of the models weights the personalized model progressively more heavily in the model-mixture analysis with increasing data.

As an example, consider learning in the Microsoft Outlook Mobile Manager (OMM) [Microsoft Corporation, 2000], a product derived from the Priorities effort. OMM provides users with a pre-built predictive model for email urgency. In the training process, numerous user-specific features are removed, so the system operates, for example, on features expected to have overall universality, such as structural aspects of email, including length of messages and the number of people on the recipients list. As a user provides cases to the system, a personal model is constructed and is continually revised with the addition of new cases. As the quantity of cases grows in the personal model, its output is weighted more heavily in combination of model outputs, until the initial predictive model is completely washed out.

Challenging areas of research include developing a better understanding of the best approaches to constructing generic models that can provide valuable, usable initial experiences with intelligent applications and services, but that allow for efficient adaptation downstream with a user's explicit training efforts or in-stream supervision. Research may lead to deeper insights about setting up systems for "ideal adaptability" given expectations about the nature of different kinds of environments, and adaptations, given the users and uses.

### 2.2 Criticality of Mixed-Initiative Capabilities

Intelligent systems with the ability to support a mix of machine and human initiatives to address problems at hand are especially critical for applications of ambient intelligence—where solutions, support, recommendations, and warnings are offered typically in stream with ongoing activities [Horvitz 2007]. There is a great opportunity for developing systems that understand how to work in a collaborative manner with users, where the system has skills in recognizing opportunities for problem solving and in understanding which aspects of problems the machine versus the person might best solve. Mixed-initiative interaction would also benefit by providing systems with the ability to infer subtleties of cognitive states of people so as guide the "if and when" of interventions. A set of principles of mixed-interactive interaction and the value of harnessing decision-theoretic principles for guiding action under uncertainty for guiding mixed-initiative interaction are presented in [Horvitz, 1999]. Work is progressing on mixed-initiative user interaction on multiple fronts, including such efforts as explorations into efficient interfaces and interactions for correcting recognition errors [Shilman et al., 2006].

### 2.3 Mental Models, Transparency, and Control

Attempts to weave machine learning and reasoning into daily life face a challenge of making the behavior of systems understandable to people. There has been little work to date on the perceptions and overall mental models of laypeople about the operation of learning and reasoning systems—and about the influence that different mental models *about* the operation of automated reasoning may have on usage, acceptance, training, and effective configuration and control of systems by people. Concerns about how automated reasoning operates and how system behaviors can be modified will likely increase as the systems move into roles and activities that people care deeply about, and as the responsibilities of systems shift from making gentle recommendations to taking higher-stakes actions in the world on behalf of people.

Understanding and addressing potential concerns with the transparency, trust, and controllability of intelligent systems is a challenging, multidimensional research area. There is opportunity to better understand peoples' mental models and to construct explanations to make the workings and conclusions of systems more transparent. In some cases, it may be useful for systems to employ explicit explanation subsystems [Suermondt, 1992] or even to modify reasoning with alternate approximations so as to enhance the understandability of the deliberation of reasoning systems [Horvitz *et al.,* 1989].

Innovations in this realm promise to allow people to better understand the basis for automated actions, to build confidence—or appropriate distrust—in the actions and policies executed in different settings, and to understand if, when, and how a system's behavior might be changed to better suit a user's wishes or expectations.

### 2.4 Privacy, Data, and Machine Intelligence

With the advent of increasingly ubiquitous sensing and reasoning, and the rise of preference and intention machines in multiple arenas, people will seek to understand and to control how personal data is being used. Developing approaches to addressing potential concerns about the privacy of data used for learning and reasoning will be important in the adoption of applications of machine intelligence. I will touch on several promising approaches to enhancing the handling of privacy in learning and reasoning systems.

**Protected sensing and personalization.** In many applications, it is feasible to perform machine learning and reasoning within a *protected shroud of privacy*, where the sensed data about activities and content of people is kept in the local control of their owners, *e.g.,* on stores within users' personal machines. For example, a geocentric intention machine with the ability to predict someone's destination while they are driving can be learned from personal GPS data that has been collected, stored, and processed locally. As another example, local analysis of potentially sensitive data, stored within a system designed to protect sensing, learning, and reasoning, can be used to enhance web search; in work on the PS prototype, a comprehensive index of a users email, documents, and web search activity is constructed locally. A large list of web results is requested from a Web search engine, and this list of, *e.g.,* several hundred results, is re-ranked with local models that consider relationships between the content of the web pages and information in a user's personal store [Teevan *et al.,* 2005].

Protected sensing and personalization could be employed in conjunction with predictive models constructed from volunteers and with third-party content. For example, a predictive model about destinations can constructed from data built from multiple volunteers, and then incorporated and used locally within a user's shroud of privacy. A large cache of advertising content might be intermittently downloaded to a user's system, but matched privately within the protective shroud of a user's machine to do local targeted advertising based on location, web pages being viewed, communications, and other content and activities.

**Learning and harnessing preferences about privacy.** The language used to define legal and organizational policies about the access and use of personal data are often limited to particular conceptions of privacy for purposes of clarity, expediency, and universal application. However, "privacy" is not a simple, nor a universal concept even within the same culture. A recent study of the sensitivities that people might have with sharing different types of information with people in different groups revealed that there are significant variations in preferences among people [Olson *et al.,* 2005]. Such results highlight the potential value of performing additional investigation of preferences about allowable uses of personal data in reasoning systems—and the application of learned insights to the design of expressive representations, interfaces, and controls that allow people to custom tailor policies for the sharing of data with other people and applications.

As a community, we need to better understand the varied and varying preferences of people with regard to the use of their personal data in learning and reasoning systems. Such understanding includes the study of potentially changing sensitivities about the use of data and the trades that people may be willing to make in cases where valuable services are offered when personal data is shared with organizations and people. Machine learning itself can be useful in probing preferences about privacy [Olson *et al.*, 2005].

**Partial revelation.** Personal data can be transformed via such processes as anonymization, summarization, abstraction, and obscuration (via such techniques as the controlled introduction of errors) in pursuit of making the sharing of the data with people, organizations—or a specific reasoning system hosted by a particular organization—more acceptable. There is an opportunity to study peoples' preferences about sharing data that has undergone different transformations. Such preferences can be coupled with analyses that provide insights about losses in the fidelity of reasoning with different transformations so as to provide guidance about the most appropriate matching of transformations to end uses.

**Restricted rights.** There is an opportunity to develop methods that annotate user data with privacy metadata which restricts the usage of the data to specified uses. Developing schemas and infrastructure to support data rights management for restricting the use of personal data would likely be a significant undertaking, and would depend on the development and widespread adoption of standards. Such a privacy infrastructure, perhaps developed as part of more comprehensive efforts to introduce standards for representing higher-level semantics on the Web, could be enabling for developing intelligent applications that employ sensitive data in a trusted manner.

## 3 Summary

I discussed trends that are leading to a nurturing environment for creating valuable applications of machine intelligence. I presented several examples of implicit and explicit applications of automated intelligence, including applications that are enabling new services and experiences. Then, I reviewed several challenges, touching on topics in the realms of learning, mixed-initiative interaction, understandability and control, and the privacy of data resources used in machine learning and reasoning. I hope that these reflections are helpful to the community of researchers pursuing the tantalizing vision of enhancing the lives of people via the fielding of applications of machine intelligence.

## Acknowledgments

I am indebted to researchers and engineers on the Adaptive Systems & Interaction team at Microsoft Research, as well as to colleagues and students at other centers who contributed insights to the projects and ideas highlighted here.

## References

[Anderson and Horvitz, 2002] Corin R. Anderson and Eric Horvitz. Web Montage: A Dynamic Personalized Start Page. *Proceedings of WWW, World Wide Web Conference*, pages 704-712, Honolulu, HI, May 2002.

[Burges et al., 2005] C. Burges, T. Shaked, E. Renshaw, A. Lazier, M. Deeds, N. Hamilton, and G. Hullender. Learning to Rank using Gradient Descent. In *Proceedings of ICML*, 2005.

[Downey *et al.*, 2007] Doug Downey, Susan T. Dumais, and Eric Horvitz. Models of Searching and Browsing: Languages, Studies, and Applications. *Proceedings of IJCAI*, Hyderabad, India, January 2007.

[Dwork et al., 2006] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor, Our Data, Ourselves: Privacy via Distributed Noise Generation, *Proceedings of Advances in Cryptology—Eurocrypt 2006*, pages 486–503, 2006.

[Fogarty *et al.*, 2005] James Fogarty, Andrew J. Ko, Htet Htet Aung, Elspeth Golden, Karen P. Tang, and Scott E. Hudson. Examining Task Engagement in Sensor-Based Statistical Models of Human Interruptibility. *Proceedings of CHI*, pp. 331-340, Portland, OR, April 2005.

[Horvitz et al. 1989] Eric Horvitz, David Heckerman, Keung Chi Ng, Bharat Nathwani, Heuristic Abstraction in the Decision-Theoretic Pathfinder System, *Proceedings of Symposium on Computer Applications in Medical Care*, Washington DC, IEEE Press: Silver Springs, MD, November 1989.

[Horvitz, 1999] E. Horvitz. Principles of Mixed-Initiative User Interfaces. *Proceedings of CHI, Conference on Human Factors in Computing Systems*, pages 159-166, Pittsburgh, PA, May 1999.

[Horvitz *et al.*, 1999] Eric Horvitz, Andy Jacobs, David Hovel. Attention-Sensitive Alerting. *Proceedings of UAI, Conference on Uncertainty and Artificial Intelligence*, pages 305-313, Stockholm, Sweden, July 1999.

[Horvitz, *et al.*, 2002] Eric Horvitz, Paul Koch, Carl.M. Kadie, and Andy Jacobs. Coordinate: Probabilistic Forecasting of Presence and Availability. *Proceedings of UAI, Conference on Uncertainty and Artificial Intelligenc*e, pages 224-233, Edmonton, Canada, July 2002.

[Horvitz *et al.* 2003] Eric Horvitz, Carl M. Kadie, Timothy Paek, and David Hovel. Models of Attention in Computing and Communications: From Principles to Applications, *Communications of the ACM*, 46(3):52-59, March 2003.

[Horvitz *et al.*, 2003b] Eric Horvitz and Johnson Apacible. Learning and Reasoning about Interruption. *Proceedings of the International Conference on Multimodal Interfaces*, pages 20-27, Vancouver, Canada October 2003.

[Horvitz *et al.*, 2003c] Eric Horvitz, Johnson Apacible, Muru Subramani., Raman Sarin., Paul Koch., J.J. Cadiz, Atilla Narin, and Yong Rui. Experiences with the Design, Fielding, and Evaluation of a Real-Time Communications Agent, *Microsoft Research Technical Report* MSR-TR-2003-98, December 2003.

[Horvitz *et al.*, 2004] Eric Horvitz, Susan T. Dumais, and Paul Koch. Learning Predictive Models of Memory Landmarks, *Proceedings of CogSci, Annual Meeting of the Cognitive Science Society*, Chicago, IL, August 2004.

[Horvitz *et al.*, 2004b] Eric Horvitz, Paul Koch, and Johnson Apacible. BusyBody: Creating and Fielding Personalized Models of the Cost of Interruption. *Proceedings of CSCW, Conference on Computer-Supported Collaborative Work*, pages 507-510, Chicago, IL, November 2004.

[Horvitz *et al.*, 2005] Eric Horvitz, Johnson Apacible, Ramin Sarin, and Lin Liao. Prediction, Expectation, and Surprise: Methods, Designs, and Study of a Deployed Traffic Forecasting Service, *Proceedings of UAI*, July 2005.

[Horvitz *et al.*, 2005] Eric Horvitz, Paul Koch, Raman Sarin, Johnson Apacible, and Muru Subramani. Bayesphone: Precomputation of Context-Sensitive Policies for Inquiry and Action in Mobile Devices. *Proceedings of User Modeling*, pages 251-260, Edinburgh, Scotland, July 2005.

[Horvitz, 2007] E. Horvitz. Reflections on Challenges and Promises of Mixed-Initiative Interaction, *AAAI Magazine* 28, 2007.

[Kapoor *et al*., 2007] Ashish Kapoor, Eric Horvitz, and Sumit Basu. Selective Supervision: Guiding Supervised Learning with Decision-Theoretic Active Learning, *Proceedings of IJCAI,* Hyderabad, India, January 2007.

[Krumm and Horvitz, 2006] John Krumm and Eric Horvitz. J. Krumm and E Horvitz. Predestination: Inferring Destinations from Partial Trajectories, *Proceedings of Ubicomp*, pages 243-260, Irvine, CA, September 2006.

[Microsoft Corporation, 2000] Outlook Mobile Manager. http://research.microsoft.com/adapt/omm/

[Olson *et al.*, 2005] Judith S. Olson, Jonathan Grudin, and Eric Horvitz: A Study of Preferences for Sharing and Privacy. *Proceedings of CHI*, pages 1985-1988, Portland, OR, April 2005.

[Shilman et al., 2006] Michael Shilman, Desney S. Tan, and Patrice Simard. CueTIP: A Mixed-Initiative Interface for Correcting Handwriting Errors. *Proceedings of Symposium on User Interface Software and Technology*, pages 323-332, Montreux, Switzerland, November 2006.

[Suermondt, 1992] Henri J. Suermondt. Explanation in Bayesian Belief Networks. *Ph.D. Dissertation*, Department of Computer Science and Medicine, Stanford University, Stanford, CA.

[Teevan *et al.* 2005] Jaime Teevan, Susan T. Dumais, Eric Horvitz. Personalizing search via automated analysis of interests and activities. *Proceedings of SIGIR, Conference on Information Retrieval*, pages 449-456, Salvador, Brazil, August 2005.