Virtual Healthcare Interaction:
Papers from the AAAI Fall Symposium (FS-09-07)

# Remote Monitoring of Activity, Location, and Exertion Levels

**Curry Guinn and Daniel Rayburn-Reeves**

University of North Carolina Wilmington
601 South College Road, Wilmington, NC 28403
guinnc@uncw.edu; djrreeves@gmail.com

**Abstract**

The purpose of this study was to develop and test a platform that would assist the Environmental Protection Agency (EPA), and the scientific community at large, in the generation of a human activity and energy expenditure database of sufficient detail to accurately predict human exposures and dose to various pollutants. The monitoring system developed is easily extendable to the collection of other health-related data. Our protocol tested the use of a digital voice recorder to collect activity/location diary data assuming it to be a less burdensome and a more reliable method than using paper and pencil diaries or hand-held computers. We expected the data to be more complete and reliable than retrospective reports (diaries filled out at the end of day) because the recorders are easy to use, the diary entries are made as the events occur, and we expected that participants would be more likely to complete the study because of the reduced burden. The data collection plan was also expected to show that the cost of the transcription of the diary can be reduced substantially by using speech and language processing to translate the digital diaries into the EPA's Comprehensive Human Activity Database (CHAD).

## Introduction

The purpose of this study was to develop and test a platform that would assist the Environmental Protection Agency (EPA), and the scientific community at large, in the generation of an activity/location/time/energy expenditure database of sufficient detail to accurately predict human exposures and dose to pollutants. Collecting daily location/activity data along with measures of pulmonary dynamics to assess human exposure to airborne pollutants has proven to be expensive, to be burdensome to the participant, and to produce unreliable data. This study tested the feasibility of enhancing existing technology to reduce respondent burden and cost and to improve the reliability of data collection.



One goal of this study was develop a protocol with highly detailed reliable information with low subject burden, two features which are generally inversely correlated. Our protocol tested the use of a digital voice recorder to collect activity/location diary data assuming it to be a less burdensome and a more reliable method than using paper and pencil diaries or hand-held computers. We expected the data to be more complete and reliable than retrospective reports (diaries filled out at the end of day) because the recorders are easy to use, the diary entries are made as the events occur, and we expected that participants would be more likely to complete the study because of the reduced burden. The data collection plan was also expected to show that the cost of the transcription of the diary can be reduced substantially by using speech and language processing to translate the digital diaries into the EPA's Comprehensive Human Activity Database (CHAD).

## Background and Previous Research

### Capturing Environmental Exposure Data

The definition of environmental exposure provided by Ott [12] and later adapted by others [9,11,15,16] is "an event that occurs when there is direct contact at a boundary between a human and the environment with a contaminant of a specific concentration for an interval of time." This definition implies that four variables must be measured to accurately characterize exposure: location (L), time (T), activity (A), and concentration (C).

With renewed EPA interest in understanding the relationship between activity and exposure, the early activity data collected as part of exposure field studies and activity pattern surveys (e.g., [6]) needed to be unified in a single representation. The EPA developed a system to store and systematically analyze the available data, the Consolidated Human Activity Database (CHAD) [10].

Relatively recent surveys to capture time activity data have been undertaken both on a state level [13] and the national level [8]. These surveys utilized the 24-hour recall method, which was a snap shot in time of individuals representative of the state or nation.

Paper-based diaries, electronic diaries, voice-recorded diaries, and observational techniques have all been used to collect data about temporal activities, spatial locations, product use, and dietary consumption of research participants [7]. Diary methods relying on recall are not highly reliable and have a relatively high respondent burden, which negatively impacts participant compliance. Observational techniques are extremely costly and burdensome. Post-study processing of diary entries is labor intensive unless simplified reporting protocols are employed and automatic processing systems are developed.

The primary task for this study was to develop or modify technology for use in the collection of longitudinal activity data. In an earlier study of CO exposures, the investigators recognized the need to automate this aspect of the data collection and pilot-tested a hand-held data entry device to be used by the participant to enter the location/activity code throughout the monitoring period [1]. This methodology utilized a programmable HP 41C calculator, where the keys were programmed to represent specific locations or activities [17]. This basic system was modified by Freeman [4] although a different micro-processor system was used and an internal clock was included in the package. More sophisticated electronics available today enable the participant to record a broader range of locations and activities by use of hand-held devices, including the Palm Pilot or Personal Digital Assistant (e.g., [2]).

## Domain of study

The utterances that serve as the basis of this research were collected as part of an Environmental Protection Agency (EPA) study that recorded voice diaries of subject activity and location information throughout the day. The subjects were instructed to enter their activities and locations into a digital voice recorder as these data changed over a seven day period. These diary entries, or utterances, were organized chronologically by subject into a database that represents the daily activities/locations of subjects in the study. A human coder then transcribed each of the audio files into a text format and classified them into appropriate activity and location semantic categories using the EPA's CHAD (Consolidated Human Activity Database) code database. The CHAD database was created by the EPA as a unified representation for activities and locations and will be the basis for the semantic categories into which each utterance will be classified [1]. For example, the semantic category for the location of the subject's own kitchen is represented by the CHAD code: **30121 Kitchen**. For example:

**Example Diary Entry: "In the kitchen about to make some eggs"**

**Location CHAD code: 30121 - Kitchen**
**Activity CHAD code: 11100 - Prepare and Clean up food.**

A snapshot of a particular subject's diary entries is given in Table I.

**Table 1: Segment from Subject 1's Diary and Corresponding CHAD Codes**

| Time | Recorded Utterance | CHAD LOCATION | CHAD ACTIVITY |
|---|---|---|---|
| 8:57 AM | in the bedroom starting housework | 30125 - Bedroom | 11200 - Indoor Chores |
| 8:59 AM | carrying clothes to the laundry room | 30128 - Utility room / Laundry room | 11410 - Wash clothes |
| 9:00 AM | the bedroom getting more clothes | 30125 - Bedroom | 11410 - Wash clothes |
| 9:05 AM | loading the washing machine in the laundry room | 30128 - Utility room / Laundry room | 11410 - Wash clothes |
| 9:06 AM | sitting down going to watch twenty minutes of Regis | 30122 - Living room / family room | 17223 - Watch TV |
| 9:23 AM | I'm going to be brushing the dog in the family room | 30122 - Living room / family room | 11800 - Care for pets/animals |
| 9:29 AM | the laundry room moving the clothes from the washer to the dryer | 30128 - Utility room / Laundry room | 11410 - Wash clothes |
| 9:34 AM | taking the dog for another walk in the rain | 30210 - Your residence, Outdoor | 11800 - Care for pets/animals |
| 9:45 AM | kitchen doing dishes | 30121 - Kitchen | 11210 - Clean-up food |
| 10:00 AM | in my office checking email | 30126 - Study / Office | 17160 - Use of computers |
| 10:10 AM | in the hallway playing ball with Fetzer | 30120 - Your residence, indoor | 11800 - Care for pets/animals |
| 10:18 AM | back to the laundry room to load and unload clothes | 30128 - Utility room / Laundry room | 11410- Wash clothes |
| 10:54 AM | pulled the sheets from the dryer I'm going to making the bed and I'm going to be switching stuff from the washer to the dryer and reloading the washing machine reloading washing machine | 30128 - Utility room / Laundry room | 11410- Wash clothes |

## Experimental Platform

Two physical devices were employed – one to capture the heart rate data; the other to capture the voice diary.

### Heart Rate Monitoring

Each subject wore a custom-built heart rate monitor with three EKG electrodes. This device recorded heart rate at 2 minute intervals. Further, if the subject's heart rate changed by more than 15 beats per minute from the previous measurement, the device would beep – a signal that the subject should make a voice diary entry.

### Voice Diaries

Subjects carried a holstered digital recorder (Sony ICD-MS1) to make their diary entries. Because of cumbersome headsets and unreliable wireless microphones, subjects were required to use the built-in microphone on the device. Before participating in the study, subjects were required to read a selection of text to create a custom profile for Dragon's Naturally Speaking speech recognition system.

During training, participants were instructed to make a voice diary entry whenever they changed location, changed activity, or if the heart rate monitor beeped.

## Experimental Results

Nine participants participated in the study for a duration of one week each in addition to the training period. A summary of some of the characteristics of these participants is given in Table 2. Across the entire data collection period for the nine participants, the average daily experiment period was 8.56 hours.

**Table 1: Subject Characteristics**

| ID | Sex | Occupation | Age | Education |
|---|---|---|---|---|
| 1 | F | Manages Internet Company | 52 | Some College |
| 2 | F | Grocery Deli Worker | 18 | Some College |
| 3 | M | Construction Worker | 35 | High School |
| 4 | F | Database Coordinator | 29 | Graduate Degree |
| 5 | F | Coordinator for Non-profit | 56 | Some College |
| 6 | M | Unemployed | 50 | High School |
| 7 | M | Retired | 76 | High School |
| 8 | M | Disabled | 62 | High School |
| 9 | M | Environment Technician | 56 | Graduate Degree |

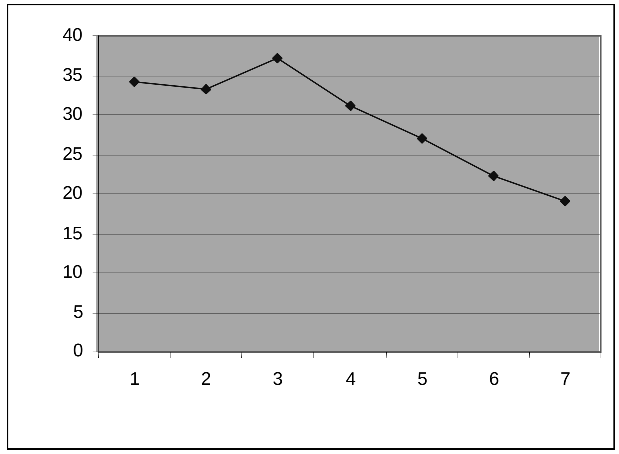



**Figure 1: Average Number of Recordings Per Day of the Study**

### Recordings per day

There were 1350 diary entries across the nine participants. The average number of recordings per participant was 29 per day. Given the average monitoring time of 8.56 hours, participants were making 3.39 recordings per hour. An interesting trend over the course of the field trial was a precipitous decline in the number of recordings as the week progressed. During the first three days of the trial, participants averaged 34.81 recordings a day. During the last 2 days of the trial, participants averaged 20.67 recordings a day. This suggests that participant fatigue of the recording process is significant. Figure 1 shows the diary reporting trend for participants over 7 days of the trial.

### Prompting based on heart rate change

In an attempt to encourage participants to report their change in activities, the heart rate monitor would prompt the participant with a tone if the participant's heart rate changed by 15 beats per minute between two recording intervals. This tone was fairly unobtrusive. During training, participants were told to record their current location and activity whenever they heard the tone. Our trial study indicates that this tone encouraged diary entries with only limited success. By analyzing the heart rate data, we can determine when a tone was issued by the monitor. By cross-referencing those times with the times of diary entries, we can determine whether an entry was made around the time of the tone. For this analysis, we assumed the participant was following protocol if they made a diary entry at any point in a 3-minute interval surrounding the time of the tone. Table 3 provides the average number of prompting tones issued per day and the percentage of times a subject made an entry corresponding to a tone.

**Table3: Heart Rate Change Indicator Tones and Subject Compliance**

| Subject | Number of Tones Per Day (Avg.) | % of Times Subject Made a Diary Entry Corresponding to a Tone |
|---|---|---|
| 1 | 22.1 | 45% |
| 2 | 41.8 | 29% |
| 3 | 32.5 | 36% |
| 4 | 33.0 | 55% |
| 5 | 33.3 | 36% |
| 6 | 15.6 | 40% |
| 7 | 32.5 | 37% |
| 8 | 26.0 | 22% |
| 9 | 22.7 | 31% |

### Diary entry quality

Of the 1350 diary entries, 133 (9.85%) contained no encodable data. The subject's inadvertent start of recording or failure to stop a previous recording was the overwhelming cause of these false entries. One disadvantage of the voice diary is that participants are given no prompts or guidelines interactively while making recordings. This problem was manifest in the data collected. In 90 entries (6.6%) the subject's activity could not be coded by a human listener. In 20 entries (1.5%) a human encoder could not determine the location. Examples of this sort of deficient diary entry might only include the location ("I'm in the kitchen") or only the activity ("I'm putting on my clothes"). It may be possible to infer the location or the activity in some cases, particularly if the previous diary entries were made in close time proximity.

### Speech recognition quality

The speech recognizer used for this project was Dragon's Naturally Speaking. Each subject was required to read a 15-minute training script (a selection from *Alice in Wonderland*). The daily recordings were made on the Sony digital recorder and the actual speech recognition was done offline. The quality of speech recognition varied widely across subjects as is illustrated in Table 4.

**Table 4: Per Word Speech Recognition Rate**

| Participant | Per Word Recognition Rate (%) |
|---|---|
| 1 | 63 |
| 2 | 54 |
| 3 | 59 |
| 4 | 61 |
| 5 | 29 |
| 6 | 17 |
| 7 | 45 |
| 8 | 49 |
| 9 | 56 |

## Statistical NLP for word analysis

The primary task is to take these spoken language utterances, such as "I am on the bus on my way to South Square Mall", and automatically select the appropriate activity and location CHAD codes. This task is commonly referred to as "text abstraction" and will be performed using statistical natural language processing techniques.

The statistical NLP technique consists of breaking the original data into a training corpus and a testing corpus. The training corpus is utilized for performing statistical analysis to build probabilities that could be used to choose the most likely semantic categories of new utterances, whereas the testing corpus is utilized provide new utterances the to test the model's classification ability. In our domain, the testing corpus was created by removing one day of data for one subject. The training set was created from the remaining data from all subjects.

The training set is analyzed to get probabilities for a naïve Bayesian classifier to classify the utterances of the testing set based on the words in each utterance. For any particular utterance in the testing set, statistics were generated based on the unigram (single words), bigram (word pairs), and trigram (word triples) probabilities and combined together to classify an utterance into a CHAD code each for activity and location [2]. For the purposes of ease in writing in this paper, when referring to unigrams, bigrams and trigrams, the phrase "word N-grams" will be used in their place.

### Bayesian classifier for N-grams

Word N-gram probabilities were determined for each utterance in the training set. As an illustration, what is the probability that the location is **Kitchen** given the presence of the word "kitchen" somewhere in the diary entry? We employ Bayes' rule:

$$P(A \mid B) = P(B \mid A) * P(A) / P(B)$$

or as an example in our particular domain:

$$P(Kitchen \mid "kitchen") = \frac{P("kitchen" \mid Kitchen) * P(Kitchen)}{P("kitchen")}$$

The formula for P("kitchen"|**Kitchen**) is computed by the percentage of times the word "kitchen" appears in utterances that have been classified as the category **Kitchen**. P(**Kitchen**) is the probability that an utterance is of the semantic category **Kitchen**, and P(“kitchen”) is the probability that “kitchen” appears in any utterance [3].

This formula can be generalized to accommodate bigrams and trigrams in the above example. Instead of a single word as an input to the function, word doubles and triples are used. Consider the following example:

$$P(Kitchen \mid "the\_kitchen") = \frac{P("the\_kitchen" \mid Kitchen) * P(Kitchen)}{P("the\_kitchen")}$$

The formulas for P(“the_kitchen”| **Kitchen**), P(**Kitchen**), and P(“the_kitchen”) are computed as they were above.

### Word-Only Score Combination

Intuitively, trigrams should be given higher weight than bigrams and bigrams higher than unigrams. For example, take the trigram “riding to work”. The activity coding of **Travel to/from…** deserves much higher precedence based on the trigram than the presence of the unigrams “riding” and “work” somewhere in the input string.

A simple linear combination is employed where, for each input and possible semantic category, the formula is:

$$\begin{aligned} Score\,(S, I) = {} & 0.2 * Unigram\ (S, I) + \\ & 0.3 * Bigram\ (S, I) + \\ & 0.5 * Trigram\ (S, I) \end{aligned}$$

The weights are chosen so they added up to one, thus the final score represented a combination of elements that are parts of a whole [4]. To determine which semantic category for each input I (I=utterance) is most likely we calculate Score(S,I) for all possible values of S (S=CHAD code for either location or activity) and chose the maximum.

## Statistical Analysis Using Surrounding Context

### Using Context to Resolve Ambiguities

Using the word-only n-grams produces 97.5% CHAD code classification accuracy for location and 96.8% classification accuracy for activity when the tested on the training set. However, the classification accuracy of the word-only system drops to 65.5% for location and 55.3% for activity CHAD codes when applied to the test corpus. This study’s primary goal is to improve the baseline accuracy by utilizing contextual information present within the diary. Specifically we assume the semantic classification of surrounding utterances is available.

The data collected from the subjects is in chronological order. Thus the sequence of activity and location CHAD codes is preserved within the data. This sequence is exploited to generate probabilities for each utterance’s CHAD code taking into account previous utterances for location and activity.

### Primary Reasoning

Common sense tells us that there is a relationship between a person’s locations and activities. One of these relationships is that your current activities and locations are related. Where you currently are and what you are currently doing is usually related to one another.

For example, if you are cooking food you are typically in the kitchen and vice versa -- it is highly unlikely that you are in the bathroom or on the sidewalk outside when you are cooking. Another of these relationships is the connection between your current location and your previous location. Typically where you are and where you were are related. Similarly, what you are doing and what you were just doing are related.

The word-only system is limited to only the words in the diary entry whereas the human encoder has access to all of the surrounding context as well as the words. A common example of a difficult entry to encode for the word-only system is a home office. An entry like “sitting at my desk” without context is ambiguous between a home office and a work office (but is distinguished in the CHAD database). Since work offices are more frequent in the corpus, the word-only system would likely select 32100 - Office building / bank / post office as opposed to 30126 - Study / Office. However, if the previous location indicated the subject was at home, this information should be factored into the analysis.

In all there were six relationships that formed the base of the context calculations:

- 3 for activity
  - What is the probability of the activity given the current location?
  - What is the probability of the activity given the previous activity?
  - What is the probability of the activity given the previous location?

- 3 for location
  - What is the probability of the location given the current activity?
  - What is the probability of the location given the previous location?
  - What is the probability of the location given the previous activity?

## Algorithms

### Context Formulas

For each contextual relationship, a formula serves as the basis of the calculations for the addition to the word-only system. The generic formulas for the six calculations are naïve Bayes probabilities of the current CHAD code given the relationships described above.

Current Location given Current Activity = $P(CLoc \mid CAct)$
Current Location given Previous Activity = $P(CLoc \mid PAct)$
Current Location given Previous Location = $P(CLoc \mid PLoc)$

A parallel structure exists for Activity CHAD codes.

### Weighting of Context Information

When calculating the word-only system, there is only one probability we were concerned with. In the case of the context and word system, we have 4 factors to combine for both location and activity. Assuming a linear combination of weights, we determine the optimal weights experimentally.

The weights are assigned to each calculation (For Activity: per word N-gram score, Current Location, Previous Activity, Previous Location ; for Location: per word N-gram score, Current Activity, Previous Location, Previous Activity) and then combined to get an overall score for a particular CHAD code classification:

$$ActivityScore = w_1 * PerWord + w_2 * PAct + w_3 * PLoc + w_4 * CLoc$$

$$Location\ Score = w_5 * PerWord + w_6 * PLoc + w_7 * PAct + w_8 * CAct$$

The weights should add up to one thus giving each calculation a probability as part of a whole. These weights were determined experimentally by trying all combinations between 0.0 and .095 with a step size of 0.05 and selecting the best scores from the training set.

## Experimental Results

### Leave-one-out testing

The total corpus consists of 1220 utterances by 9 different subjects. Due the relatively small size of the data set, leave-one-out testing was utilized to remove any test set bias from the results. There were 42 days of subject utterances within the full data set, thus there were 42 testing set/training set tests run in this experiment.

When the word-only system was applied to the leave-one-out data sets, the results were as follows: location CHAD code classification accuracy was 65.5%, activity CHAD code classification accuracy was 55.3%.

In the following example presented in Table 5, the subject's diary entry "in the office at the computer" actually corresponds to the subject being on the computer in her home office. The word-only analysis strongly favors the subject being in an office building because the phrases "in the office" and "computer" are strongly biased towards being in an office building in the training corpus. However, the previous diary entry was encoded with the location being **Living room/family room**. This additional context information inhibits coding the current utterances location as **Office building/bank/post office** and increases the likelihood of **Study or Home Office**.

**Table 5: Using Context Information Corrects Semantic Interpretation**

| Diary Entry: "in the office at the computer"<br>Correct Location: **Study or Home Office**<br>Previous Location: **Living room / family room** | |
|---|---|
| **Top 3 Location Word-only Choices (w/ probability)** | **Top 3 Location Context + word Choices (w/ prob.)** |
| 0.904 - Office building/bank/post office | 0.502 - Study / Office |
| 0.217 - Public building/library/museum /theater | 0.349 - Office building/bank/post office |
| 0.053 - Public garage / parking | 0.212 - Public building/library/museum /theater |

### Adding Contextual Information

When contextual information is combined with the original word-only calculations to the entire test set, the results scores are summarized in Table 6.

**Table 6: Word-only Analysis vs. Word + Context Information**

| | Word-only | Context & word | Improvement |
|---|---|---|---|
| **Locations** | 65.5% | 76.0% | 16% |
| **Activities** | 55.3% | 66.9% | 21% |

### Discussion

Looking only at the words in the diary entry, we achieved an average accuracy of 65.5% for locations and 55.3% for activities. When the context data was combined with the word calculations, there was a marked improvement. Applying the context system to the test sets, the location accuracy improved to an average of 76.0% an improvement of 16% and the activity accuracy improved to an average of 66.9% an improvement of 21%.

With the leave-one out testing weights, there was no single weight configuration that provided optimal results for all of the data sets. Further research would have to be done to see if there is a pattern, range or some other general method for generalizing weights for all leave-one out data sets. The average weights that performed well are presented in Table 7.

**Table 7: Relative Weighting of Words and Context Information**

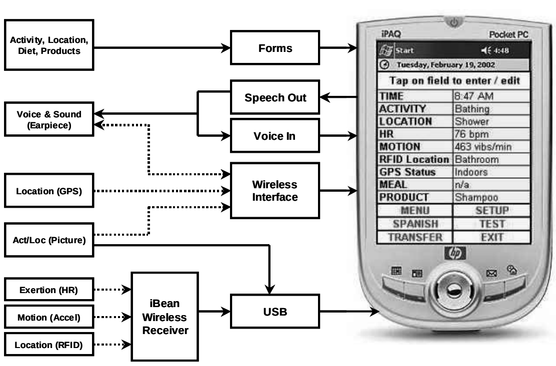
| Activity | | | Location | | |
|---|---|---|---|---|---|
| **Per-word** | $w_1$ | 0.35 | **Per-word** | $w_5$ | 0.29 |
| **Previous Activity** | $w_2$ | 0.18 | **Previous Location** | $w_6$ | 0.15 |
| **Previous Location** | $w_3$ | 0.20 | **Previous Activity** | $w_7$ | 0.27 |
| **Current Location** | $w_4$ | 0.27 | **Current Activity** | $w_8$ | 0.29 |

## Future Work

One drawback of the system described above is that data is captured passively by the system with no immediate feedback to the user. In practice, the only time the system communicated to the subject was when the system issued a beep if the subject's heart rate changed significantly. This beep was a signal to the subject to make a diary entry.

A better solution is to create a dialog between the subject and the system with more elaborate prompting and immediate error-correcting dialogs when the system is unable to process the user's diary entry. To create this system, we are currently building an input device with the PocketPC technology using a wireless mic/headset, built-in GPS, a pen-based diary, and peripherals for exertion and motion data gathering (Figure 2).

**Figure 2: Time, Activity, Location, Exertion Data Gathering Platform**




This system will also fuse data from multiple input sources and use discrepancies to trigger dialogs with the user. For instance, the system will have constant access to Global Positioning System data which includes not just coordinates but also speed. If the voice diary entry indicates that the person is at home but the GPS reports the position is away from home and traveling at 55 mph, then the system can prompt the user to re-enter the voice diary entry. In some situations, the system could prompt with very specific questions with expected Yes/No answers – a response which has a high likelihood of correct speech recognition.

## References


[1] Akland, G.G., Harwell, T.D., Johnson, T.R., and Whitmore, R.W. "Measuring human exposure to carbon monoxide in Washington, D.C., and Denver, Colorado during the winter of 1982-1983." *Environ. Sci. Technol.* 19:911-918. 1985.

[2] Brauer, M., Hirtle, R.D., Hall, A.C., and Yip, T.R. Monitoring personal fine particle exposure with a particle counter. *J. Expos. Anal. Environ. Epidem.* 9:228-236. 1999.

[3] Burton, R. R. *Semantic Grammar: a Technique for Efficient Language Understanding in Limited Domains*. Doctoral Thesis. UMI Order Number: AAI7629270. 1976.

[4] Freeman, N.C.G., Waldman, J.M., Lioy, P.J. Design and evaluation of a location and activity log used for assessing personal exposure to air pollutants. *J Expos Anal Environ Epi*, 1: 327-338. 1991.

[5] Hipp, D.R., *Design and development of spoken natural-language dialog parsing systems*. Ph.D. thesis, Duke University, available as Technical Report, CS-1993-15, 1992.

[6] Johnson, T. *A study of human activity in Cincinnati, Ohio.* Contractor report for Electric Power Research Institute, Palo Alto, CA, 1987.

[7] Johnson, T., Long, T., and Ollison, W. *A survey of prospective technologies for collecting human activity data.* Contractor report for the American Chemistry Council. 2001.

[8] Klepeis, N.E., Nelson, W.C., Tsang, A.M., and Robinson, J.P. The national human activity pattern survey (nhaps): data collection methodology and selected results.) *J. Expos. Anal. Environ. Epidem.* 1998.

[9] Lioy, P.J. Assessing total human exposure to contaminants. *Environ. Sci. Technol*. 1990. 24:938-945.

[10] McCurdy, T., Glen, G., Smith,L., and Lakkadi,Y. The National Exposure Research Laboratory's Consolidated Human Activity Database. *J. Exp. Anal. Environ. Epidem*. 10:566-578, 2000.

[11] National Academy of Sciences (NAS). Human exposure assessment for airborne pollutants; advances and opportunities. National Academy of Sciences, Washington, DC, 1991.

[12] Ott, W.R. "Concepts of human exposure to air pollution." Environ. Intl. 7:179-196. 1982.

[13] Robinson, J.P., Wiley, J.A., Piazza, T., Garrett, K., and Cirksena, K. *Activity Patterns of California Residents and Their Implications for Potential Exposure to Pollution*. Sacramento, CA: California Air Resources Board (CARB-A6-177-33), 1989.

[14] RTI, Longitudinal activity data for selected susceptible population subgroups. Final Report for EPA Contract No. 68-D-99-012 (Task Order 0012). 2001.

[15] U.S. Environmental Protection Agency (USEPA). Guidelines for exposure assessment. *Federal Register,* 57(104):22888-22938, 1992.

[16] Zartarian, V.G., Ott, W.R., and Duan, N. A quantitative definition of exposure and related concepts. *J. Expos. Anal. Environ. Epidem*. 7:411-437, 1997.

[17] Zelon, H.S. Capture of activity pattern data during environmental monitoring, in: T.H. Starks (ed.) *Proceedings of the Research Planning Conference on Human Activity Patterns.* Las Vegas NV: U.S. Environmental Protection Agency, pp. 8-1 to 8-12, 1989.