Association for Information Systems
AIS Electronic Library (AISeL)

MCIS 2011 Proceedings

Mediterranean Conference on Information Systems (MCIS)

2011

HCI MARKERS: A CONCEPTUAL FRAMEWORK FOR USING HUMAN-COMPUTER INTERACTION DATA TO DETECT DISEASE PROCESSES

Yoram M. Kalman
The Open University of Israel, yoramka@openu.ac.il

HCI MARKERS: A CONCEPTUAL FRAMEWORK FOR USING HUMAN-COMPUTER INTERACTION DATA TO DETECT DISEASE PROCESSES

Kalman, Yoram M, The Open University of Israel, 1 University Road, Raanana, 43107, Israel, yoramka@openu.ac.il

Abstract

HCI markers are those signals created during human-computer interaction (HCI) which might provide information about the cognitive, mental, psychological or physiological state of the user. This conceptual paper presents the concept of HCI markers and makes the analogy between these markers and biomarkers, which are used in medicine to identify the possibility of unusual physiological conditions such as cancer. It then suggests a list of variables that have been reported in the literature and which might potentially fill the role of HCI markers. It suggests a course of study to systematically identify these markers, and discusses the risks associated with the HCI marker approach. It emphasizes the extensive multidisciplinary and interdisciplinary work that is required to achieve the goal of making HCI markers a part of the toolkit of modern medicine.

Keywords: Human-computer interaction, Health, Markers, Alzheimer's disease, Emotion, Personality, Privacy

1. INTRODUCTION

Users who interact with information systems create a constant stream of data. The data are created through keyboards and pointing devices (computer mouse, touch screen, etc.), as well as via other sensors such as cameras, microphones and GPS receivers. This paper suggests that this stream of data has value which is beyond the immediate utilitarian purposes of the interaction, and that analyzing this stream could provide information about the health of the users. It presents research findings which link specific variables related to human-computer interaction (HCI) with specific cognitive, mental and psychological states of the users, and suggests that these and other yet unstudied variables can be used in a manner similar to the use of biomarkers.

Biomarkers are "physical, functional or biochemical indicators of physiological or disease processes. These key indicators can provide vital information in determining disease prognosis, in predicting of response to therapies, adverse events and drug interactions, and in establishing baseline risk." (Future Medicine, 2011). Biomarkers have been linked to various conditions such as cancer, Alzheimer's disease (AD), metabolic disorders, stress, and cardiovascular disease (Diamandis, 2010; Foundation for the National Institutes of Health, 2011; Gerszten & Wang, 2008; Hellhammer, Wüst, & Kudielka, 2009). Like other binary classification measures, the challenge in biomarker research is to optimize both the sensitivity of the assay, as well as its specificity (e.g. Nickolas, et al., 2008; Perkins & Schisterman, 2006).

In this conceptual paper I show that several variables related to HCI have already been linked to specific cognitive, mental and psychological variables. I suggest that these preliminary findings imply that it would be possible to identify HCI related variables which are associated with important health-related conditions. I then suggest several research trajectories to explore these possibilities, and discuss the pitfalls which need to be taken into account in such research and in the implementation of its findings.

2. HCI MARKERS

2.1 HCI markers and neurodegeneration

A host of variables that might serve as HCI markers have been reported in the literature so far. Many of them have to do with neurodegeneration, since neurodegeneration often influences areas in the brain which are involved in language production. For example, in a paper titled "cognitive archaeology" Garrard (2009) discusses several retrospective studies of language decline in written texts and in transcribed spoken language samples. These texts were produced by people who were later diagnosed with slowly progressive dementias. These studies are described below.

In a longitudinal study of members of the School Sisters of Notre Dame religious congregation, Snowdon et al. (1996) analyzed autobiographies written by the nuns when they joined the order in the 1930's and 1940's. They compared the linguistic characteristics of each early autobiography with the present day performance of the nun in cognitive tests. The analysis showed that low idea density and low grammatical complexity in those early autobiographies were associated with low cognitive test scores dozens of years later. More specific to neurodegeneration, they also showed that of the 14 nuns who died by the time the study was conducted, all five who had neuropathologically confirmed AD had low idea density in their early biographies, compared with none of those without neuropathologically confirmed AD. When autobiographies of 11 more nuns from other convents who were neuropathologically evaluated were similarly examined, similar results were obtained. In total, in the cohort of 25 nuns, 90% of those with AD showed low idea density in texts written in their early life, compared with 13% in those without AD.

In a study of the impact of AD on the written language of a single author whose AD was confirmed post-mortem (Garrard, 2008; Garrard, Maloney, Hodges, & Patterson, 2005), a comparison was made between early works by the celebrated writer Iris Murdoch (IM), and her latest novel, Jackson's Dilemma. The novels were not edited. Based on the fact that AD reduces syntactic complexity and lexical richness, three of IM's books were analyzed: an early one written in 1954 (Under the Net), The Sea, The Sea, written at the peak of her career (1978), and Jackson's Dilemma, written during the early stages of her AD. The analysis detected several quantitative measures marking the impact of early AD on lexical properties of the written language: the vocabulary became less varied (decreased type-to-token ratio) and included more frequent words.

Another such retrospective study was carried out by Brian Butterworth, who studied the language used by former US President Ronald Reagan during his debates with Walter Mondale in 1984, and comparing it to the language used by Reagan four years earlier (Forbes-McKay & Venneri, 2005; Garrard, 2009). He detected an increased rate of errors in content and in syntax, as well as abnormally long word-finding pauses. Ten years later, in 1994, Reagan was diagnosed with AD. Attempts to perform a similar retrospective study of speeches made by another politician, the British Prime Minister Harold Wilson, also showed marked differences in his language prior to his sudden resignation in early 1976. The reason for the resignation was probably Wilson's awareness of cognitive changes associated with the preclinical phase of a progressive degenerative dementia, probably AD (Garrard, 2009; Gomez, 2010).

2.2 HCI markers and emotion

HCI markers can also be used to identify changing emotional states. A good example is work performed by Pennebaker, Mehl and colleagues in relation to the shock experienced by Americans following the September 11, 2001 attacks (Cohn, Mehl, & Pennebaker, 2004; Liehr, Mehl, Summers, & Pennebaker, 2004). The markers in these cases were based on the vocabulary used by the study population. Their vocabulary was analyzed using the text analysis program Linguistic Inquiry and Word Count (LIWC: Pennebaker, Mehl, & Niederhoffer, 2003). LIWC checks the vocabulary used in a document against an internal dictionary of words and word stems which have been assigned to specific linguistic categories.

In the Cohn et al. study, over 70,000 blog entries spanning the September 11 events were analyzed for words indicating emotional positivity, cognitive processing, social orientation and psychological distancing, as well as for words indicating preoccupation with the September 11 attacks. The results show the expression of more negative emotions immediately following the attacks, as well as increased social and cognitive engagement, and increased psychological distancing. The effect on psychological distancing was shown to last longer, while other effects, such as emotional positivity, returned to baseline within a short period of time.

2.3 Additional HCI markers

Language can also reflect other variables such as personality traits, and level of interpersonal trust (Kalman, Scissors, & Gergle, 2010; Mairesse, Walker, Mehl, & Moore, 2007). There are preliminary findings that identify individual differences between users based on their keystroke behavior (Epp, Lippold, & Mandryk, 2011; Khanna & Sasikumar, 2010; Rao, 2005). As these and other studies show (Castellano, Kessous, & Caridakis, 2008), more variables can provide more precise information, and thus it is valuable to move beyond signals created solely via the keyboard, and include signals that can be collected via microphones, cameras and pointing devices. This allows the inclusion of variables such as facial expressions, vocal pitch and intensity, variables which can provide detailed information that can be analyzed to teach us about fleeting, as well as more stable emotional states of the users (Cowie, et al., 2001).

This section reviewed published work on only several human traits such as personality, cognitive ability and emotion. These traits are, of themselves, of medical importance, and as we have seen in the examples relating to AD, can reflect individual changes of significant medical importance. However, the significance of these examples goes beyond the obvious suggestion already made by some of the researchers, to use these markers to monitor for signs of changes such as cognitive deterioration. The significance of the examples lies in their demonstration that signals created during HCI can reflect fundamental attributes of the users. This suggests that such variables might also reflect many other attributes that have not yet been studied, and that some of these might be of medical relevance. What might these be?

If user emotion and personality are reflected in the variables created during HCI, it might be possible to identify HCI markers for mental illness. Preliminary steps in this direction have been described in a study of language in messages posted to a bipolar disorder chatroom (Kramer, Fussell, & Setlock, 2004). Or, if an illness such as glaucoma leads to a deterioration in eye-hand coordination (Kotecha, O'Leary, Melmoth, Grant, & Crabb, 2009), could this decreased coordination influence the usage of pointing devices such as a computer mouse? Also, could measuring mouse movements lead to an earlier diagnosis of tremor that might be associated with Parkinson’s disease or other medical conditions such as essential tremor (Shahed & Jankovic, 2007)? These are only a few speculative suggestions, but the potential of linking HCI markers with medically important conditions is great.

3 IDENTIFYING HCI MARKERS: TOP DOWN AND BOTTOM UP

The previous section described several HCI marker candidates, based on research carried out in the last two decades. The challenge this section tackles is what systematic research could be suggested that would identify informative HCI markers which are reliable and valid enough for medical purposes.

One bottom up approach is based on large cohorts of essentially healthy participants who will be recruited into the study, and agree that their interactions with their various digital devices (desktops, laptops and mobile devices such as cell phones and tablet computers) be recorded over a period of years. These longitudinal records will then be periodically analyzed and compared with the changes in the medical status of the study participants. Correlations between changes in specific markers and the development of medical conditions will be identified and compared to unaffected controls, and their reliability and validity evaluated. Such bottom up approaches are usually based on extensive data mining techniques to analyze very large datasets (e.g. Ghazavi & Liao, 2008; Klee, 2008; Rhodes, et al., 2004).

A top down approach would examine specific medical conditions and try to predict, based on the etiology and symptoms of each condition, which HCI markers might be correlated with it. Then, archival records of online interactions by patients who have this condition will be examined for these markers. Such archival records could include personal correspondence such as email or IM, as well as more public documents such as blogs and micro-blog postings. The work on AD described above (e.g. Garrard, 2009; Gomez, 2010) is a good example of this approach in which the researchers examine linguistic characteristics which are associated with AD, and examine ways to detect changes in these characteristics in archived datasets.

The main drawback of the bottom up approach suggested here is that it is lengthy, and requires a large cohort of participants. Consequently, it will require extensive financial and human resources, and usable findings will emerge slowly. The main drawbacks of the top down approach suggested here is that archival records might be difficult to acquire, that these records include only a subset of the data created during HCI (for example typing speed and mouse movements are not captured in archived email messages), and that they will still require a large cohort of controls to verify the specificity of the putative marker. Given these drawbacks, a recommended approach would combine both ad-hoc top down work, with a more systematic long term bottom up longitudinal study. The longitudinal study might also include participants who already have medical conditions of interest. The interim findings of the bottom up research can be used to fine tune the top down studies, as well as to provide reference data in lieu of separate control cohorts for each of the top down studies.

4 DATA COLLECTION

This section discusses the opportunities and challenges faced by researchers who wish to collect the data required for the extraction of HCI markers. This discussion focuses on the collection of digital HCI data for the purpose of monitoring users for unusual trends in HCI markers that have already been identified. Similar principles apply to collecting data from research subjects for the purpose of identifying novel markers.

Unlike most current research on HCI markers which focused on retrospective collection of archived data, i.e. of the final product (blog post, novel, transcript of IM-based conversation), our goal in the future should be to collect all data related to the production of that final product, as well as any additional information created while interacting with the digital device. For example, blog posts or email messages contain little information beyond the actual content of the message. On the other hand, a log of the keystrokes that were used to create the post or the message (Sullivan & Lindgren, 2006) can teach us about the chronemic variables associated with the production of the message (e.g. typing speed, pauses during the typing) as well as about revisions of wording, spelling or page layout.

Information about the production process that led to the messages is important not only as a source of HCI markers. It also provides information about the context of the interaction that is taking place. For example, keystroke logging can provide information about the usage of spelling and grammar checkers; GPS sensors can provide information about whether the message was produced at home or at work, or while walking down a busy street; and, microphones can monitor the audio ambience in which the user is immersed. All of these contextual elements can assist in interpreting the data stream more effectively.

Which devices can be used for monitoring HCI markers? In principle, any device with which users interact and which produces digital output that could be collected and archived is a candidate. The obvious candidates are workstations and portable computers, as well as mobile devices such as cell phones, smart phones, and tablet computers. Less obvious candidates are mobile and wearable devices such as audio and video players (e.g. iPod), e-readers (e.g. Kindle), gaming platforms (e.g. Kinect) head-mounted displays, or electronic textiles (Barfield & Caudell, 2001). Early lessons can be learned about the ability of such systems to collect and analyze data about the user in the emerging field of wearable physiological monitoring devices (e.g. Otto, Milenkovic, Sanders, & Jovanov, 2006; Pandian, et al., 2008; Wu, et al., 2008). Unlike these wearable physiological monitoring devices which are focused on the moment to moment monitoring of vital signs, it is expected that, at least in the short and medium term, HCI markers will only be used for detecting long term trends that impact health on a scale of weeks, months and years.

Although the collection of data is carried out at the level of the individual, the findings could be aggregated to study groups of people or even whole populations, in order to increase our understanding of the development of specific conditions in individuals as well as population-wide processes (e.g. Brownstein, Freifeld, & Madoff, 2009; Cohn, et al., 2004; Liehr, et al., 2004).

5 USING THE RESULTS

In what way could HCI markers contribute to healthcare? The first and most obvious use is for diagnostic purposes: to identify a change in behaviour that is sensed through HCI and which might indicate the early stages of illness. Since the markers can be monitored in an unobtrusive manner, and since the analysis of the data is not expected to be expensive, these markers can complement periodic medical checkups and possibly even be evaluated in conjunction with the medical data which is used in those check-up appointments. It is possible to envision a future at which HCI data analysis is as routine as a blood or urine test, and as simple as a blood pressure measurement.

Early detection is one of the more attractive advantages of HCI markers, since the data stream is created on a daily basis, and can be analyzed as often as necessary. One criticism of early detection arises in the case of conditions which, at present, cannot be treated (e.g. AD) or conditions in which the impact of early detection on eventual mortality, morbity and quality of life is controversial (e.g. prostate cancer: Bartsch, et al., 2008; Wolf, et al.). Despite these controversies regarding the value of population-wide screening, detection of pre-symptomatic and early stage patients is critical for research purposes and for the evaluation of experimental treatments. Without the ability to detect the earliest possible stages of the progression of a condition, it is significantly more difficult to evaluate treatments that might halt or slow the progression of a disease. Since some sectors of the population interact with computers extensively (Pew Research Center, 2010) these sectors produce a very large amount of data on a daily basis. These extensive datasets could turn out to be highly sensitive to the behavioural changes that signal the onset of conditions which are currently impossible or difficult to detect, or which require to proactively ask for a test to be carried out.

There are many other possible uses of HCI markers. Like other biomarkers, they can also assist in predicting the response to therapies and in monitoring these responses, as well as in establishing baseline risk (Future Medicine, 2011). One relatively unusual property of HCI markers is related to the fact that HCI data can be archived for a long period of time. For example, many users of web-based email services such as gmail and Yahoo! Mail can use gigabytes of storage to archive their emails. This means that HCI markers can be used as a “time machine” which allows collecting pre-diagnostic behavioural data after a diagnosis of an illness occurs. Such a retrospective study can inform medical personnel about the early and sometimes even “pre-symptomatic” stages of the condition. This information could be used for research purposes, as well as to better understand the patient’s own condition and its evolution.

It is interesting to note that datasets collected for the purpose of identifying HCI markers can also be used for other HCI studies. Specifically, they can be used to allow better customization of interfaces to user needs, and specifically to the needs of users with relevant health conditions.

6 LIMITATIONS AND RISKS

HCI markers have the potential to provide rich information similar to that provided by ubiquitous standardized medical tests such as blood or urine tests. It is relatively inexpensive to collect, and since all data are already digitized, there is no need for expensive transcription or digitization costs. The development of effective automated analysis tools will facilitate the creation of an end-to-end digital process that requires minimal human intervention, only at junctures that involve decision-making. This analysis can either be carried out in real time (for an example of real time analysis of such data see Leshed, et al., 2009), or at predetermined intervals, in batch mode. The ease of collection and analysis of ubiquitous digital data is also at the heart of the key risks associated with the proposal to use HCI markers for medical purposes. These include risks associated with reliability and validity, and risks associated with privacy and security.

Like other biomarkers and health-related variables, HCI markers too have limited reliability and validity (Diamandis, 2010), which have to be carefully assessed before the markers are implemented. It is not expected that it would be difficult to identify HCI-related variables that correlate with the progression of specific conditions. The challenge would be to ascertain the extent to which these changes are specific to the condition. For example, we expect many HCI-related variables to change with aging, and the challenge would be to distinguish between these naturally occurring changes, and changes associated with a specific illness. Variability in the HCI-related variables will result not only from long term processes such as aging. As already demonstrated, these variables are influenced by ephemeral states such as moods, and the question is whether the naturally occurring variability will not mask the changes we expect HCI markers to detect. Another challenge is specificity: like body temperature or blood pressure, some of the important HCI markers will probably not be specific to one condition, but rather flag a circumstance which requires further evaluation. These reliability and validity issues are an important risk of the HCI-marker approach since, unlike lab tests which require highly specialized equipment, analysis of digital data created during HCI can be performed by almost any advanced PC user, and freely available online applications can even widen the circle of potential (ab)users of these markers. The risks associated with wide and essentially unsupervised analysis of such markers are exemplified by the significant ethical and medical challenges caused by the rising availability of direct-to-consumer genetic testing (Hudson, Javitt, Burke, & Byers, 2007; Patch, Sequeiros, & Cornel, 2009).

The collection of HCI data requires the monitoring and collection of large amounts of private information. Careful attention needs to be paid to ensure that this does not negatively impact the privacy and security of study participants and of future adopters of this approach. Here too we can learn from experience gained in the field of genetic analysis, where the ability to sequence whole genomes opened a host of new ethical challenges for researchers, as well as for practitioners (McGuire, Caulfield, & Cho, 2008). Ethical and legal issues will need to be carefully considered when collecting, storing, analyzing, and using the analysis results. These will need to be translated into clear informed consent procedures that will ensure that participants understand the risks associated, and consider those risks acceptable. More specifically, since the same variables that characterize the individual and that are HCI marker candidates could also serve as biometric variables that identify the individual to secured systems (e.g. Rao, 2005), special attention needs to be paid to this intersection of privacy and security (e.g. Kerr, Steeves, & Lucock, 2009; McGuire, et al., 2008).

Another privacy related concern raised by the prospect of HCI markers is the ability to analyze data produced by users without their permission. If blog posts, email messages, status updates, and other forms of public and semi-public online communication would turn out to reveal medically sensitive information, then it might be possible to take archives of such messages and learn about the cognitive, mental, psychological or physiological state of the user. Do people have a right to prevent others from performing this analysis or from using the results? Some aspects of these dilemmas are similar to those encountered in regards to rights associated with the extraction and analysis of DNA samples (Berson, 2009; Michael, 2010; Nerko, 2008), but in essence this is a new and complex question which should be considered before the capability to perform such analysis becomes a widespread reality.

7 CONCLUSION

This conceptual paper presents and discusses the concept of HCI markers, analogous to biomarkers used in medical fields. After a review of studies that have demonstrated associations between these markers and changes in emotional, cognitive and psychological states, it is suggested that HCI markers could be an unobtrusive, low cost, and powerful tool to detect and track conditions of medical importance. The potential power and relative simplicity and low cost of this tool also raise significant concerns about misuse and abuse of this approach, and the threats to privacy and to security are discussed. Some lessons can be learned from the past in regards to similar issues related to other markers (e.g. genetic markers) and other illnesses (e.g. cancer and AD), but there are also significant novel risks associated with this approach that have not been encountered in the past.

The promise of HCI markers is great, but it will only be fully realized through carefully controlled longitudinal studies in the general population, as well as studies of these markers in carefully selected populations with diagnosed emotional, cognitive, psychological or physiological conditions, and by diligent attention to the social and psychological risks associated with the collection, analysis, and decision making processes that will be developed. These will require multidisciplinary and interdisciplinary approaches that include, in addition to the already multidisciplinary approach of HCI, professionals from the life and medical sciences, from social and behavioural sciences, and from the humanities.

8 ACKNOWLEDGEMENT

I would like to thank Darren Gergle and Kathleen Geraghty for helpful discussions.

9 REFERENCES

Barfield, W., & Caudell, T. (2001). *Fundamentals of wearable computers and augumented reality*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bartsch, G., Horninger, W., Klocker, H., Pelzer, A., Bektic, J., Oberaigner, W., et al. (2008). Tyrol Prostate Cancer Demonstration Project: early detection, treatment, outcome, incidence and mortality. *BJU International, 101*(7), 809-816.

Berson, S. B. (2009). Debating DNA Collection. *National Institute of Justice Journal, 264*, 9-13.

Brownstein, J. S., Freifeld, C. C., & Madoff, L. C. (2009). Digital disease detection—harnessing the web for public health surveillance. *New England Journal of Medicine, 360*(21), 2153-2157.

Castellano, G., Kessous, L., & Caridakis, G. (2008). Emotion recognition through multiple modalities: face, body gesture, speech. *Affect and emotion in human-computer interaction*, 92-103.

Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science, 15*(10), 687.

Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., et al. (2001). Emotion recognition in human-computer interaction. *Signal Processing Magazine, IEEE, 18*(1), 32-80.

Diamandis, E. P. (2010). Cancer Biomarkers: Can We Turn Recent Failures into Success? *Journal of the National Cancer Institute, 102*(19), 1462-1467. doi: 10.1093/jnci/djq306

Epp, C., Lippold, M., & Mandryk, R. L. (2011). *Identifying Emotional States Using Keystroke Dynamics*. Paper to be presented at the CHI 2011, Vancouver, BC, Canada.

Forbes-McKay, K. E., & Venneri, A. (2005). Detecting subtle spontaneous language decline in early Alzheimer's disease with a picture description task. *Neurological Sciences, 26*(4), 243-254.

Foundation for the National Institutes of Health. (2011). The Biomarkers Consortium Retrieved April 24, 2011, from http://www.biomarkersconsortium.org

Future Medicine. (2011). Biomarkers in Medicine: Aims and Scope Retrieved April 24, 2011, from http://www.futuremedicine.com/page/journal/bmm/aims.jsp

Garrard, P. (2008). The Iris Murdoch text analysis project and its importance to the study of authorship and Alzheimer's disease. *The Iris Murdoch Review, 1*(1), 14-17.

Garrard, P. (2009). Cognitive archaeology: Uses, methods, and results. *Journal of Neurolinguistics, 22*(3), 250-265.

Garrard, P., Maloney, L. M., Hodges, J. R., & Patterson, K. (2005). The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author. *Brain, 128*(2), 250.

Gerszten, R. E., & Wang, T. J. (2008). The search for new cardiovascular biomarkers. *Nature, 451*(7181), 949-952.

Ghazavi, S. N., & Liao, T. W. (2008). Medical data mining by fuzzy modeling with selected features. *Artificial Intelligence in Medicine, 43*(3), 195-206.

Gomez, P. C. (2010). Analysing linguistic decline in early-stage Alzheimer's disease: a corpus-based approach. *A Mosaic of Corpus Linguistics: Selected Approaches*, 165.

Hellhammer, D. H., Wüst, S., & Kudielka, B. M. (2009). Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology, 34*(2), 163-171.

Hudson, K., Javitt, G., Burke, W., & Byers, P. (2007). ASHG statement on direct-to-consumer genetic testing in the United States. *American Journal of Human Genetics, 81*(3), 635.

Kalman, Y. M., Scissors, L., & Gergle, D. (2010, September 12-14). *Chronemic aspects of chat, and their relationship to trust in a virtual team.* Paper presented at the Mediterranean Conference on Information Systems, Tel-Aviv, Israel.

Kerr, I., Steeves, V. M., & Lucock, C. (2009). *Lessons from the identity trail: anonymity, privacy and identity in a networked society*: Oxford University Press, USA.

Khanna, P., & Sasikumar, M. (2010). Recognising Emotions from Keyboard Stroke Pattern. *International Journal of Computer Applications, 11*(9), 24-28.

Klee, E. W. (2008). Data Mining for Biomarker Development: A Review of Tissue Specificity Analysis. *Clinics in Laboratory Medicine, 28*(1), 127-143.

Kotecha, A., O'Leary, N., Melmoth, D., Grant, S., & Crabb, D. P. (2009). The Functional Consequences of Glaucoma for Eye–Hand Coordination. *Investigative ophthalmology & visual science, 50*(1), 203.

Kramer, A. D. I., Fussell, S. R., & Setlock, L. D. (2004). *Text analysis as a tool for analyzing conversation in online support groups*.

Leshed, G., Perez, D., Hancock, J. T., Cosley, D., Birnholtz, J., Lee, S., et al. (2009). *Visualizing real-time language-based feedback on teamwork behavior in computer-mediated groups.* Paper presented at the CHI Boston, MA.

Liehr, P., Mehl, M. R., Summers, L. C., & Pennebaker, J. W. (2004). Connecting with others in the midst of stressful upheaval on September 11, 2001. *Applied Nursing Research, 17*(1), 2-9.

Mairesse, F., Walker, M. A., Mehl, M. R., & Moore, R. K. (2007). Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text. *Journal of Artificial Intelligence Research, 30*, 457-500.

McGuire, A. L., Caulfield, T., & Cho, M. K. (2008). Research ethics and the challenge of whole-genome sequencing. *Nature Reviews Genetics, 9*(2), 152-156.

Michael, K. (2010). *The Legal, Social and Ethical Controversy of the Collection and Storage of Fingerprint Profiles and DNA Samples in Forensic Science.* Paper presented at the IEEE International Symposium on Technology and Society, Wollongong, NSW, Australia.

Nerko, C. J. (2008). Assessing Fourth Amendment Challenges to DNA Extraction Statutes after Samson v. California. *Fordham Law Review, 77*, 917.

Nickolas, T. L., O'Rourke, M. J., Yang, J., Sise, M. E., Canetta, P. A., Barasch, N., et al. (2008). Sensitivity and specificity of a single emergency department measurement of urinary neutrophil gelatinase–associated lipocalin for diagnosing acute kidney injury. *Annals of internal medicine, 148*(11), 810.

Otto, C., Milenkovic, A., Sanders, C., & Jovanov, E. (2006). System architecture of a wireless body area sensor network for ubiquitous health monitoring. *Journal of Mobile Multimedia, 1*(4), 307-326.

Pandian, P. S., Mohanavelu, K., Safeer, K. P., Kotresh, T. M., Shakunthala, D. T., Gopal, P., et al. (2008). Smart Vest: Wearable multi-parameter remote physiological monitoring system. *Medical engineering & physics, 30*(4), 466-477.

Patch, C., Sequeiros, J., & Cornel, M. C. (2009). Genetic horoscopes: is it all in the genes? Points for regulatory control of direct-to-consumer genetic testing. *European Journal of Human Genetics, 17*(7), 857-859.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology*, 547-578.

Perkins, N. J., & Schisterman, E. F. (2006). The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. *American journal of epidemiology, 163*(7), 670.

Pew Research Center. (2010). Online Activities, Daily. *Internet & American Life Project* Retrieved May 2, 2011, from http://www.pewinternet.org/Static-Pages/Trend-Data/Online-Activities-Daily.aspx

Rao, B. (2005). *Continuous keystroke biometric system.* Masters of Science in Media Arts and Technology Project paper, University of California, Santa Barbara, , Santa Barbara. Retrieved from http://ilabtest.cs.ucsb.edu/publications/RaoMS.pdf

Rhodes, D. R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., et al. (2004). ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia (New York, NY), 6*(1), 1.

Shahed, J., & Jankovic, J. (2007). Exploring the relationship between essential tremor and Parkinson's disease. *Parkinsonism & related disorders, 13*(2), 67-76.

Snowdon, D. A., Kemper, S. J., Mortimer, J. A., Greiner, L. H., Wekstein, D. R., & Markesbery, W. R. (1996). Linguistic ability in early life and cognitive function and Alzheimer's disease in late life. *JAMA: The Journal of the American Medical Association, 275*(7), 528-532.

Sullivan, K. P. H., & Lindgren, E. (2006). *Computer keystroke logging and writing: Methods and applications*: Emerald Group Publishing.

Wolf, A. M. D., Wender, R. C., Etzioni, R. B., Thompson, I. M., D'Amico, A. V., Volk, R. J., et al. American Cancer Society Guideline for the Early Detection of Prostate Cancer: Update 2010. *CA: A Cancer Journal for Clinicians, 60*(2), 70-98. doi: 10.3322/caac.20066

Wu, W. H., Bui, A. A. T., Batalin, M. A., Au, L. K., Binney, J. D., & Kaiser, W. J. (2008). MEDIC: Medical embedded device for individualized care. *Artificial Intelligence in Medicine, 42*(2), 137-152.