A Data Framework to Understand the Lived Context for Dementia Caregiver Empowerment

Marta Belay, Charles Henry, Daran Wynn, Tonya Smith-Jackson
Department of Industrial and Systems Engineering, NC A &T State University, Greensboro, NC 27411
mwbelay@aggies.ncat.edu, tlsmithj@ncat.edu

Abstract

Agitation in dementia patients is characterized by several features, such as physical and verbally aggressive and non-aggressive behaviors. Such behaviors affect not only the patients, but also their caregivers' quality of life. The onset of agitated behaviors can be unpredictable and can also be influenced by environmental factors, which introduce challenges to caregivers when caring people with dementia (PWD). The purpose of this study is to analyze multiple forms of qualitative and quantitative data obtained through behavioral and environmental sensors. Data about body gestures, activity and task sequences, ambient light, sound and temperature will be obtained. Caregiver logs and medical history from nurses and psychiatrists are the sources of qualitative data. Data framework will be used to collect, structure, extract, analyze, interpret and integrate various formats and large amount of data. This approach helps to conceptualize the lived context of PWD. The information discovered will be used to generate trained models to identify the patterns of agitation associated with the environmental factors. It will also be used to develop a monitoring and dashboard system so caregivers and healthcare providers can understand and avoid environmental triggers. The research outcome will provide cost effective technology to reduce or prevent agitation in dementia.

1. Introduction

Dementia is a general term, which describes conditions characterized by decline in memory or cognitive function that affects a person's ability to perform day-to-day activities (Alzheimer's Association, 2014). The occurrence rate of all types of dementia among individuals older than 71 was 13.9% in 2002 (Plassman et al., 2007). This rate corresponds to 3.4 million individuals in the USA. The prevalence rate of dementia has been found to increase with age from 5% of those aged between 71 and 79 years to 37.4% of those aged 90 and older (Plassman et al., 2007). The most common cause of dementia is Alzheimer's disease. In the United States, an estimated 5.2 million people have Alzheimer's disease and it is estimated that in every 67 seconds someone develops the disease. By the mid-century, the occurrence is estimated to be every 33 seconds (Alzheimer's Association, 2014). According to the World Health Organization (WHO) statistics, about 35.6 million people are affected by dementia worldwide, and Alzheimer's disease contributes to 60-70% of the cases (WHO, 2012).

Agitation is a common and challenging consequence of dementia, which occurs in 90% of the patients (Colombo et al., 2007). Various stages of dementia require different sets of skills from caregivers, and most caregivers do not have training in possible interventions. This results in stress and increased caregiver burden and also leads to institutionalization of patients in long term care facilities (Steinberg et al., 2008). In addition, it incurs higher economic cost to provide the necessary care for a person with dementia (PWD). The cost of dementia care in 2010 was estimated to be between $157 billion and $215 billion by a nationwide study (Hurd et al., 2013). In 2013, the estimated economic value of care provided by unpaid caregivers was $220.2 billion. Similarly, aggregate cost of care provided with payment was $214 billion (Alzheimer's Association, 2014).

Empowering caregivers to reduce stress and agitation in PWD will have positive impacts on the PWD, the caregiver, and the associated cost of care can be reduced. The following is a case scenario describing the experience of a caregiver attending her mother from Alzheimer's association webpage (www.alz.org).

"*I've been the primary caregiver for my mother with dementia/Alzheimer for the past nine years. She's 86 and is fading away by inches and by bits and pieces. It is so unbelievably cruel and torturous to watch someone who was an excellent teacher and active lover of life be whittled away by this hideous disease a tiny bit at a time. I'm convinced she contracted it through hormone replacement therapy, which she had for too long and past the age of 75. I really don't know how to convey how horrible this is for her and for me. She has suffered more than we can ever know, both physically and mentally. I have given 20 percent of my life to caring for her 24/7. Predictably, my life has received no* *attention at all. I have no husband, no family, no career, no retirement, and no plans for the future. I've had to endure my own personal heartaches in silence, including losing several beloved pets over the years, losing relatives and my own battle with skin cancer. Everything is secondary when you are a caregiver. Your life is forfeited, and because this battle cannot be won, you will ultimately fail. There is simply no way to put a good face on this experience*."

Such stories are common among caregivers. Caregivers of PWD have a 50% chance of experiencing depression due to the stressors they experience with the changed behavior, unpredictability, reduced cognitive abilities and role changes (Schulz et al., 1995). Caregivers with depression have increased morbidity and mortality (Pruchno and Potashnik, 1989), and PWD in these dyads have shorter times before institutionalization (Schulz et, al., 1999). Institutionalization may be linked to a more rapid psychological decline, since the individual is placed in an unfamiliar environment at a critical period and becomes cared for by individuals they do not know.

To address this significant challenge, a research team, comprised of investigators at NC A&T State University, University of Virginia and the Carilion Center for Healthy Aging, has focused on the goal to identify engineering-based interventions to increase caregiver empowerment through the use of tools to predict and minimize agitation episodes among PWD. The envisioned system, Behavioral and Environmental Sensing and Intervention (BESI), is a complex cyber-socio-physical system that incorporates technologies, social dyads and contexts. In other words, the Cyber-socio-physical system is consisted of three subsystems which comprise of various components. This complex system will be used to acquire multiple forms of descriptive data to build a knowledge base of the ecosystem surrounding agitation. Data will be analyzed to understand the lived context of a PWD and to develop a model that can be used to predict agitation events associated with the environmental conditions. A monitoring system which recognizes agitation epochs will be developed to send real time notification for caregivers. Secured web-based interface monitoring system will be used to display the sensor data for health care providers, caregivers and other authorized users. The web-interface will be refined with input from nurses, caregivers, and health informatics to ensure it is user friendly and easily interpreted. Sensor data will be grouped by category such as physical agitation, temperature and noise level, and other environmental stimuli. Users can further navigate the interface to view data from individual sensors.

As a result, caregivers can intervene on the PWD and the environment before agitation escalates. BESI will be an empowering tool for caregivers of PWD with cost effective solution. Yet, the challenge of BESI lies in the immensity of the data; where data forms, types, sources, and scales vary extensively. This paper describes the frameworks that serve as taxonomies and ontologies to assist our research team to plan, collect, extract, analyze, and interpret the multiple data streams from the BESI project. As the research is at its early stage, the conceptual data framework which facilitates the data collection and analysis, and which also forms the basis for the advancement of the technology is presented in this paper.

2. Literature Review

In modern data-intensive science, more consideration has been given to the challenges of handling massive data formats and volumes. Considering the data ecosystem as a whole is very essential to truly address the challenges of very diverse multidisciplinary data. Understanding complex system problems involving heterogeneous and diverse interdisciplinary research data requires mixed data integration and analysis (Parsons et al., 2011). A conceptual data framework can be used to map the relationships and dependencies among various scientific data sources, types of data produced and used, and curation activities associated with the data (Cragin et al., 2010). Conceptual mapping of data frameworks can also be used to reduce qualitative data, analyze themes and interconnections in the data (Onwuegbuzie et al., 2009). Data frameworks are helpful to identify types of data to be collected and data analysis techniques to be used (Parsons et al., 2011). They also serve as aids to develop new methods of analysis.

The three V's (volume, variety and velocity) are characteristics of big data. Volume refers to the large amount of data, variety refers to different types of data and velocity stands for the rate of data accumulation (Berman, 2013). The greatest benefit of big data is the ability to link seemingly different disciplines for the purpose of developing and testing hypotheses that cannot be approached within a single knowledge domain. With few exceptions, big data is ordinarily analyzed in incremental steps; the data are extracted, reviewed, reduced, normalized, transformed, visualized, interpreted and re-analyzed with different methods (Bari et al., 2014). Big data has many implications for patients, healthcare providers, researchers and other healthcare constituents. It will also impact how these players engage with the healthcare ecosystem, especially when external data, regionalization, mobility and social networking are involved (Murdoch and Detsky, 2013).

Data mining, knowledge extraction, information discovery, information harvesting and data pattern processing are some of the names used in the past to refer to the process of finding useful patterns in data (Fayyad et al., 1996), also known as knowledge discovery. Fayyad et al. (1996) define knowledge discovery as a series of activities for making sense of data. They distinguish data mining as a specific step in the knowledge discovery process which focuses on the application of certain algorithms to extract useful information (knowledge). In contrast to these distinct views of knowledge discovery and data mining, Peng et al. (2008) use combined process of data mining knowledge discovery (DMKD). They define DMKD as extraction of useful information (knowledge) from data and this extraction is achieved by learning new methods and techniques. These methods and techniques are used in the pre-processing and post processing of data, specifically for discovering previously unknown patterns and building predictive models from the data (Peng et al., 2008; Maimon et al., 2010).

Previous works which focus on monitoring agitation behaviors were reviewed. Bankole et al. (2012) conducted a study to explore the ability of a custom inertial wireless body sensor network (BSN) to detect and quantify agitation. The initial study was focused on validating the BSN. The research work consisted of data collection on selected subjects at different times of the day. From assessment of the pilot results, it was concluded that the BSN was a valid measure of agitation. The ability of the BSN for continuous and real-time monitoring was also examined (Bankole et al., 2011).

In summary, a data framework is used to plan, structure, and organize different data formats and large amounts of data for data analysis and data integration in multidisciplinary research. It helps to integrate, process, visualize, and present data in a meaningful way. Knowledge discovery processes are implemented to prepare, select and cleanse data. Proper interpretations of mined data from the research domain are possible using these processes. Constructing an integrated and interactive data framework with the application of knowledge discovery and data mining will provide a map of mixed analytical landscape for multidisciplinary researchers. This data framework can also facilitate research team communication, collaboration and the development of shared mental models. Most importantly, data frameworks support reasoned action when analyzing data. If frameworks are organized and agreed upon ahead of time while researchers are focusing on the primary research questions and objectives, the analytical processes and reasoning from the data will be more aligned with the line of inquiry established by the problem to be addressed and the research goals. In this way, research integrity is maintained.

3. Purpose of the Research

The purpose of this research is to develop and implement a data framework for the research and design team to apply data structuring and analysis on complex systems. The investigators are developing a cyber-socio-physical system to assist caregivers and providers in the management of agitation in dementia. The cyber-socio-physical system is a complex system based on its characteristics – interrelatedness, autonomous components, and dynamic.

The study to be conducted will use a remote ethnographic approach to collect data about the physical agitation of a PWD and the natural living environments of the PWD and caregivers. This is achieved by making use of different sensors on the patient as well as the surrounding environment. Body-worn sensors are placed on the PWD, which capture the movement of the patient at multiple parts of the body to detect different stages of physical agitation. Environmental acoustic sensors are installed to capture information about ambient noise and speech features. Light and temperature sensors are used to measure ambient environmental conditions. Additional set of motion sensors are installed near doorways to detect movement from one room to another. The sensor networks, wireless devices and laptop-based stations (physical structures), algorithms and computations constitute the cyber subsystem.

Different subjective measures are used to recognize the agitation events, frequency, type, and stress level experienced by the caregivers. These measures will help to quantify agitated behaviors and their impact on the PWD and caregiver separately as well as on the dyad as a unit. PWD, caregiver-healthcare providers, the PWD-Caregiver dyad, patient's family and friends make up the social subsystem. Agitation is influenced by a number of environmental factors such as ambient temperature, sound and light level, social density etc. It is important to track knowledge of this environment which makes up the physical subsystem to minimize the occurrence of agitation events in the patients.

The problem space addressed by this research is three fold:

1. The volume and variety of the data requires an organizing data framework that guides input, structuring, and analysis of the various forms of data
2. The complexity of the system (cyber-socio-physical) requires a data framework to organize team members' integrated mental models as the system is developed from concept to final prototype
3. The data framework is needed to facilitate the data to design translation process to achieve the final outcomes to benefit caregivers and PWD

4. Data Framework Development Process and Results

Developing a conceptual framework for a specific study incorporates a system of concepts, assumptions, expectation, beliefs and theories that support the research

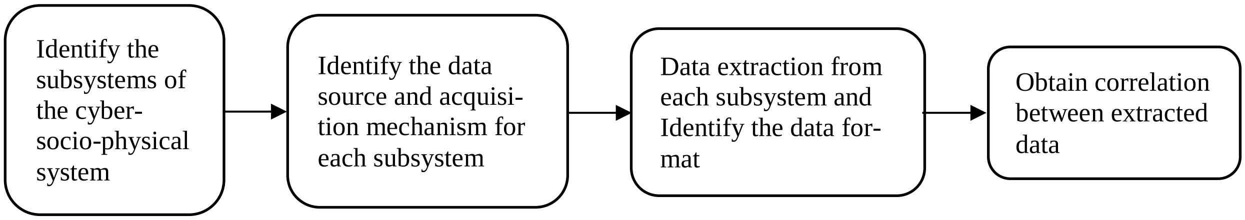

Figure 1. Flow chart that shows processes followed to generate data framework

(Wang et al., 1995). Idea association can be regarded as the catalyst that facilitates the interaction among researchers and design participants. By linking the researchers' and designers long term memory internally and previous participant knowledge externally, diverse design ideas can be generated (Lai and Chang, 2006).In the BESI project, a team that consists of multidisciplinary experts from computer and electrical engineering, human factors and ethnography, geriatric psychiatry, and nursing conducted a brainstorming session to enhance their previous knowledge about the cyber-socio-physical system with additional innovative perspectives. Individual ideas were linked with greater technical depth to generate the following flow chart. The flow chart shows the basic steps followed to generate the integrated and inclusive data framework.

Figure 1 demonstrates the data framework of the BESI project. This data framework accounts for the interactions of the components of the cyber-socio-physical subsystems. Cyber-socio-physical systems are comprised of three subsystems. The cyber subsystem consists of the inertial body-worn sensors, environmental sensors, wireless Bluetooth devices and computers. The sensor stream provides continuous data about body motions and ambient living space conditions. Similarly, the wireless Bluetooth gives information about the location of person in a house in different times of the day and night. A base workstation communicates with all the sensors and wireless devices. Data extractions from the sensory devices are done by applying different signal processing algorithms. The extracted data will have different format such as binary, continuous, ratio.

The social subsystem encompasses the dementia patient, caregivers, nurses, patient's family and friends, health care providers. Interviews, caregiver diaries, assessment batteries are used to collect data. The collected data provides information about behavioral pathology in dementia patients, cognitive level, aggressive and non-aggressive agitation symptoms, dementia stage, functional capacity, sleep quality, quality of life for the caregivers and patients, etc. Content analysis and score calculations are applied to filter useful data from the collected information. It is important to understand the various levels of social subsystems within the BESI system. There are individual social subsystems, dyadic social sub-systems (i.e., PWD-caregiver; caregiver-healthcare provider), and group-level subsystems (more than two individuals).

The Physical/environmental subsystem consists of environmental conditions that surround the dementia patients. Temperature, sound and light intensity, physical movement and speech features represent the physical subsystem. Ambient conditions and gross movement data are gathered using environmental and door way sensors. Various formats (i.e. relative frequencies of codes, ratio, categorical, binary) of the desired information are extracted from the collected data using algorithms.

The researchers can use the following data framework to plan and organize their data collection, extraction, and analysis activities. For instance, to monitor the behavior of a PWD and their environment, body-worn and environmental sensors are used. These sensors continuously provide data about the physical movement of the patient and ambient conditions in the room. However, sensory raw data is often difficult to understand and interpret, especially when sensory data comes from multiple sensors. It should be noted that the sensor data is collected every second and when this frequency of data collection is repeated for multiple sensors, the researchers would be facing challenge of handling and interpreting large amounts of data. Therefore, researchers should divide the complex system into subsystems and then apply the data framework to extract and interpret the data from each subsystem independently.

In addition, the data interpreted from each subsystem should be integrated and correlated with each other to have knowledge of the whole system behavior. The sensory data from the cyber and physical subsystems, for instance, should be converted to meaningful form to identify patterns of movement which allow categorization of the patient's behavior and the environmental conditions respectively. This can be achieved by simultaneous collection and separate interpretation of the data from both subsystems. The interpreted data are then integrated to identify the environmental condition which contributes to the agitation.

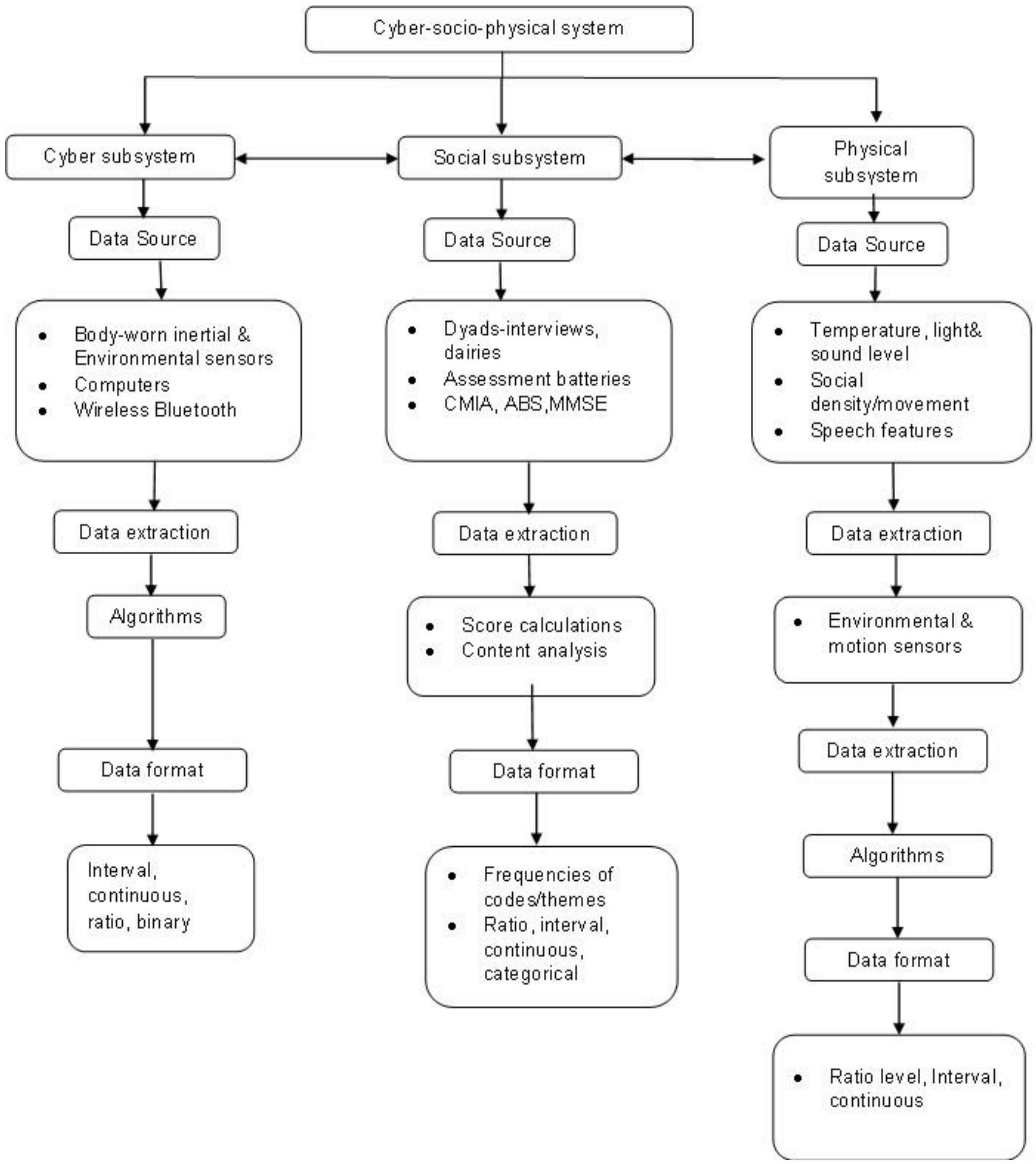

Figure 2. Cyber-socio-physical system data framework

5. Discussion and Conclusion

A complete data framework provides researchers the advantage of dealing with complex systems in several ways. It enables identification of subsystems of a complex system and helps to identify individual subsystem data sources and data acquisition mechanisms. It provides a platform towards extraction of useful information from the different subsystems. It helps to correlate the data from the subsystems which are useful for understanding of the entire system which in our research is definition of the lived context of PWD's. The project will have very diverse and large data sets to be collected and analyzed (Table 1). This table was generated through mock-up data set that is currently under development. The time scales, formats, sources, and numerical scales will differ. The large volume of data, frequency of data collection, and variety of the data which come from multiple sensors demand data framework that guides inputs, structuring, extraction, and analysis. This framework will also evolve more as the research team proceeds through the different project phases.

Verification and validation of basic BESI sensing and environmental assessment will be done in a controlled set-up and in the homes of PWD's. Reliability of body-worn

Table 1. Example of sensor data with time stamps: BSN and TM represent body worn and environmental sensor ID.

Sensor ID	Time stamps	Desired information	Data type
BSN 001	11/06/2015 04:45:34	Pre-agitation, agitation event	Interval, continuous
BSN 005	11/06/2015 04:50:56	Agitation period, post agitation	Ratio level, continuous
TM 008	11/06/2015 17:15:25	Temperature level	Ratio level, continuous
MS 007	11/06/2015 23:12:55	Social density	Binary
Caregiver diary	11/06/2015 07:05/05	Agitation frequency, level, caregiver self-reports of psychosocial variables	Relative frequency of themes, codes, quantitative ratings (qualitative themes, continuous scale ratings)
Assessment battery	12/07/2015 08:15:25	Level of cognition, functional capacity	Ratio, interval, continuous

sensors is validated based on caregiver diaries and assessment batteries of agitation events. Data from caregivers will be obtained through tablet diaries with structured prompts and a time-stamped report. This data is simultaneously collected with body worn-sensor data. Constant comparison, narrative and content analysis are used to analyze qualitative data from caregivers, whereas one or more algorithms and statistical techniques are used to analyze data from sensor streams. The results of quantitative analysis of both quantitative and qualitative data are combined at the interpretation level to validate the accuracy of sensor activities. However, each data set remains analytically separate from each other. Therefore, to plan the data collection and analysis processes at the intersection of the two subsystems, it is important to use BESI data framework.

Data mining knowledge discovery (DMKD) approaches will be carried out to extract useful information (knowledge) from the each of the subsystems. Useful information from sensor streams will be analyzed to identify agitation epochs (pre-agitation, agitation, post-agitation) and to model associated environmental conditions, which lead to slow or rapid perseverance of agitation episodes. Besides, sensor data will be combined with subjective data to support an integrated observation of the social context of an agitation event. Finally, the analyzed data and the model developed will be used to inform design of monitoring system in which caregivers and health care providers can get just in time notification about the agitation epochs and triggers.

Data frameworks in multi-disciplinary research facilitate team communication and exchange of information in the process of understanding and analyzing complex systems. It enhances knowledge and shared experience about which data analysis techniques can be integrated, how and why they are combined. In this regard, the BESI data framework will provide an excellent perspective to the importance of multi-disciplinary research collaborations to address critical societal needs.

Acknowledgement

This research is jointly funded by National Science Foundation and National Institute of Health under one Award number IIS-1418622.

References

Alzheimer’s Association. Alzheimer’s disease Facts and Figures. Includes a special report on women and Alzheimer diseases. *Alzheimer’s & Dementia* Volume 10, Issue 2, 2014.

Alzheimer’s association, Muffet’s Story
http://www.alz.org/living_with_alzheimers_15512.asp

Bankole, A., Anderson, M., Smith-Jackson, T., Knight, A., Oh, K., Brantley, J.S., Barth A.T., Lach, J. 2012. Validation of Non-Invasive Body Sensor Network Technology in the Detection of Agitation in Dementia. *American Journal of Alzheimer’s Disease & Other Dementias* 27(5):346-354.

Bankole, A., Anderson, M., Knight, A., Oh, K., Smith-Jackson, T., Hanson, M. A., Barth A.T., Lach, J. 2011. Continuous, Non-Invasive Assessment of Agitation in Dementia Using Inertial Body Sensors, *Wireless Health*, San Diego, USA.

Bari, A., Chaouchi, M., Jung, T. 2014 *Predictive analytics for dummies* New Jersey John Wiley & Sons, Inc.

Berman, J.J. 2013 *Principles of big data: Preparing, sharing, and analyzing complex information* Waltham, MA: Elsevier Inc.

Colombo, M., Vitali, S., Cairati, M., Vaccaro, R., Andreoni, G., Guaita, A. 2007. Behavioral and psychotic symptoms of dementia (BPSD) improvements in a special care unit: a factor analysis. *Archives of Gerontology and Geriatrics* 44 (Supplement 1): 113-120.

Cragin, M. H., Palmer, C. L., Chao, T. C. 2010. Relating data practices, types, and curation functions: An empirically derived framework. *Proceedings of the American Society for Information Science and Technology*. 47(1): 1–2. Pittsburgh, USA

Daley, B. J, 2004. Using concept maps in qualitative research. Concept Maps: Theory, Methodology, and Technology. *Proceeding of the first international conference on Concept Mapping*. Pamplona, Spain.

Fayyad, U., Pitaetsky-shapiro, G., Smyth, P. 1996 Knowledge Discovery and Data Mining: Towards a Unifying Framework. *KDD-96 Proceedings*.

Hurd, M. D., Martorell, P., Delavande, A., Mullen, KJ., Langa, K. M. 2013. Monetary Costs of Dementia in the United States. *The New England Journal of Medicine* 368: 1326-1334.

Lai, I. C. and Chang T. W. 2006. A distributed linking system for supporting idea association during the conceptual design stage. *Design studies* Vol 27, Issue 6, Pages 685–710.

Maimon, O. and Rokach, L. 2010. *Data Mining and Knowledge Discovery Handbook Second Edition*. Springer.

Murdoch, T. B. and Detsky, A. S. 2013. The Inevitable Application of Big Data to Health Care. *The Journal of the American Medical Association* Vol 309, No. 13.

Onwuegbuzie, A. J., Dickinson, W. B., Leech, N. L., Zoran, A. G. 2009. A Qualitative Framework for Collecting and Analyzing Data in Focus Group Research. *International journal of qualitative methods* 8(3): 1-21.

Parsons, M. A., Godøy, Ø., LeDrew, E., De Bruin, T. F., Danis, B., Tomlinson, S., Carlson, D.2011. A conceptual framework for managing very diverse data for complex interdisciplinary science. *Journal of Information Science* 37(6): 555–569.

Plassman, B. L, Langa, K. M, Fisher, G. G., Heeringa, S. G, Weir, D. R., Ofstedal, M. B., Burke, J. R., Hurd, M. D., Potter, G. G., Rodgers, W. L., Steffens, D.C., Willis, R.J., Wallaceg, R.B. 2007. Prevalence of Dementia in the United States: The Aging Demographics, and Memory Study *Neuroepidemiology*, 29(1-2): 125-132.

Peng, Y., Kou, G., Shi, Y., Chen, Z. 2008. A descriptive framework for the field of data mining and knowledge discovery. *International Journal of Information Technology & Decision Making*. Vol. 7, No. 4: 639–682.

Pruchno, R.A. and Potashnik, S.L. 1989. Caregiving Spouses. Physical and mental health in perspective. *Journal of the American Geriatric Society* 37(8):697-705.

Schulz, R., O'Brien, A. T., Bookwala, J., Fleissner, K. 1995. Psychiatric and physical morbidity effects of dementia caregiving: prevalence, correlates, and causes. *Gerontologist* 35(6):771-779.

Schulz, R. and Beach, S.R. 1999. Caregiving as a risk factor for mortality: the Caregiver Health Effects Study. *Journal of the American Medical Association* 282(23):2215-2219.

Steinberg, M., Shao, H., Zandi, P., Lyketsos, C. G., Welsh-Bohmer, K. A., Norton, M. C., Breitner, J. C. S., Steffens, D. C., Tschanz, J. T., and Cache County Investigators. 2008. Point and 5-year period prevalence of neuropsychiatric symptoms in dementia: the Cache County Study. *International Journal of Geriatric Psychiatry* 23(2): 170–177.

Trochim, W.M.K. 2006. Research methods knowledge base. Retrieved on March 20, 2015. http://www.anatomyfacts.com/research/researchmethodsknowledgebase.pdf

Wang, Y .R., Storey, C.V., and Firth, P.C. 1995. A Framework for Analysis of Data Quality Research. *IEEE Transactions on knowledge and data engineering* Vol. 7, no. 4.

World Health Organization (WHO), Media Center: Dementia. Fact sheet No.362, Apr 2012.