Week 1: Chapter 1: Draft 0x01

Nate Bachmeier

DIS-9991: Components of a Dissertation

January 16, 2022

Northcentral University

# Chapter 1: Introduction

A demographic timebomb will create significant pressure on the global health care system because people live longer, have fewer children, and medical costs continue to increase (Piggott, 2016; Stone, 2017). When patients cannot afford the required care, the quality decreases, or social programs must fund the difference. Demographic specialists predict that by 2050 nearly “80% of the global elderly population will be from low- to middle-income countries (Muhsin, Munyogwa, Kibusi, & Seif, 2020, p. 1).” Economic constraints within those countries will limit the effectiveness of their welfare programs and the availability of adequate services. Additionally, over one billion globally have a limiting disability that requires additional support (Morris, 2008). Medical facilities need mechanisms to defuse the situation by reducing costs and deferring the transition to assisted living centers.

Inversely, the explosive growth across IoT, Cloud, Big Data, and Mobile (ICBM) continuously decreases costs and enables new opportunities. These technologies have the potential to revolutionize the health care and wellbeing industries. Academic and commercial vendors are continuously delivering innovations across these domains. However, mainstream offerings primarily focus on measuring simple body metrics (Koreshoff, Robertson, Leong, 2013). While these products provide incremental value, they do not move the needle. Nearly eight years later, the industry myopically drives toward wearable IoT devices (Tun et al., 2021). Researchers concentrating on these areas make sense due to the low barrier to entry. Though, that same ease is commoditizing the products selection and stifling creativity.

Technology within special needs and elderly care settings has unique challenges and requirements (Ferati, Kurti, Vogel, & Raufi, 2016). These persons need unobtrusive systems that continuously monitor and respond to their behaviors. Specific vendors utilize voice-enabled Personal Digital Assistants (PDA) (e.g., Amazon Alexa) to effectively set reminders and record activities (Tan et al., 2020). However, it becomes challenging to globalize these voice-specific technologies to assist non-native speakers and individuals with vocal disorders.

Assisted living facilities use trained nurses to mitigate these issues. Having a human inspect the patient visually is an effective but expensive tool. The median compensation rate for registered nurses is $75,330 annually ($36.22 per hour) (US Bureau of Labor Statistics, 2020). Due to the high cost, few patients have private nurses and receive fractional supervision. In contrast, video-centric monitoring and Human Activity Recognition (HAR) apply to a diverse population. When a person falls or drinks a glass of water, their skeleton moves in predictable ways, enabling AI/ML processes to respond through CPS systems. Businesses could deliver these capabilities economically and consistently across global markets, ultimately improving the quality of care at lower costs.

However, ethical concerns and privacy issues prevent researchers from collecting data at scale. Image the complexity that small-to-medium businesses face between vetting volunteers and ensuring diversity across participants. There are also budgetary considerations to deploying IP cameras and other CPS in numerous households. These challenges prevent quality research from occurring and improve patients’ quality of care. Instead, processes must exist to simulate these interactions and iterate toward more sophisticated systems.

## Statement of the Problem

Senior citizens live longer than ever before and want to defer moving into nursing homes until later in life. Transitioning into elderly care comes as a double edge sword. On the one hand, nurses can provide 24-hour supervision. This assistance could mean the difference between life and death (e.g., during a fall). However, on the other hand, the medical services are prohibitively expensive, nearing $90,000 annually (Tan et al., 2020). Additionally, these medical facilities lack the personalization available within one’s home. Further, specific individuals with diseases like dementia and Alzheimer’s require even greater levels of attention.

Traditionally, addressing the situation requires increasing human capital, such as more traveling nurses or family member oversight (Westergren et al., 2021). However, this solution increases health care costs and collects limited patient health data. In addition, these infrequent visits might miss critical issues, especially with those most reluctant to relocate. Modern solutions must bridge the differentiation between remaining in the home and still receiving the attentiveness typically found in assisted living facilities (Tan et al., 2020). When this gap narrows, the patient can remain in familiar settings for more prolonged periods. That situation has numerous psychological benefits (e.g., higher morale) and economic impacts (e.g., deferring private health care costs). For instance, patients with memory impairment might forget to empty the dishwasher, take medication, or bathe regularly. Medical facilities can address these challenges through real-time video monitoring services that analyze patients’ actions and recommend care.

## Purpose of the Study

This constructive research demonstrates a simulation procedure for collecting human data in private and sensitive contexts. It aims to show this capability by combining various artifacts under the real-world scenario of elderly and special needs care. These existing artifacts include resources spanning MoCap databases, physics simulators, and AI/ML CV algorithms. While this specific project examines elderly care, the implications are generalizable to other scenarios. Those scenarios encompass childcare (e.g., babysitting), school safety systems, and virtual office secretary situations, to name a few. Beyond privacy, the approach applies to high-risk health and safety research. For example, it would be challenging to set numerous apartments ablaze to assess an evacuation procedure. However, actors can perform animation sequences within virtual environments and enable researchers to observe those behaviors.

Building these capabilities and verifying them at scale is challenging. First, the research must find patients willing to share continuous in-home video streams. In addition to the privacy concerns, it would be difficult for others to reproduce the findings. Third, validating the solution against numerous home layouts requires significant effort. Lastly, purchasing and configuring hardware components is prohibitively expensive in terms of time and money (Elloumi et al., 2020; Das et al., 2019). This project mitigates these issues through a virtual world simulation process.

The Robot Operating System (ROS) is a framework for writing robot software (Stanford Artificial Intelligence Laboratory et al., 2018). It exposes features for rapidly designing complex cyber-physical interactions through a Message Passing Interface (MPI). The meta-operating system integrates with physics engines (e.g., Gazebo) and machine learning platforms (e.g., OpenAI Gym). Additionally, developers can package these tools into containerized workloads and leverage public cloud services (e.g., Amazon Web Services and Microsoft Azure). The cloud enables researchers to validate their designs in numerous world permutations efficiently and economically. Furthermore, the software can test situations that are not practical or feasible within the physical world (e.g., set the kitchen ablaze). Together, these different technologies culminate into an elegant solution that monitors, predicts, and responds in real-time to patient needs.

This dissertation leverages these tools to implement an intelligent home simulation environment. Next, it will populate the virtual home with ROS devices and sensors to observe and respond to ROS actors (patients). The actors will perform animation sequences based on motion-capture records. Lastly, the researcher will assess the observations and responses against the labeled data. While this specific test scenario focuses on elderly care, the solution is broadly applicable to any Cyber-Physical simulation.

## Introduction to Research Methodology and Design

Design-science is a standard methodology for researching Information Technology (IT) problems. Hevner et al. (2004) propose a collection of guidelines for implementing this methodology (see Table 1). There are three phases to implementing this process. First, the researcher(s) must identify a domain-specific challenge. Next, that researcher creates artifacts that study this phenomenon. Third, those artifacts assess the topic and communicate answers to the research questions.

Table 1: Design-science Guidenlines (Hevner et al. 2004)

|  |  |
| --- | --- |
| Guideline | Description |
| Design as an Artifact | Design-science research must produce a viable artifact in the form of a construct, a model, a method, or an instantiation. |
| Problem Relevance | Design-science research aims to develop technology-based solutions to important and relevant business problems. |
| Design Evaluation | A design artifact's utility, quality, and efficacy must rigorously demonstrate well-executed evaluation methods. |
| Research Contributions | Effective design-science research must provide transparent and verifiable contributions to design artifacts, foundations, and/or design methodologies. |
| Research Rigor | Design-science research relies on rigorous methods to construct and evaluate the design artifact. |
| Design as a Search Process | The search for a compelling artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment. |
| Communication of Research | Design-science research must be presented effectively both to technology-oriented and management-oriented audiences. |

This dissertation employs this methodology to improve special needs and elderly care with AI/ML and CV applications. Scalability, security, and privacy challenges prohibit studying this topic through traditional means. People are generally unwilling to undergo 24/7 video monitoring and disclose their most intimate conversations in the name of science. Future research needs to address those concerns. Meanwhile, this effort provisions industry-standard physics simulation environments to examine those interactions. Next, this project creates virtual devices (e.g., IP cameras) to extract a subject’s behavior and respond accordingly. Third, a data telemetry collection pipeline will assess the performance of virtual devices within a simulated world.

## Research Questions

### RQ1

What mechanisms are best suited for extracting the subject’s *intent* when dealing with noisy video stream data? Noise enters the processing pipeline from numerous situations, such as out-of-focus images and the subject’s distance to the camera.

### RQ2

What affordances do Cyber-Physical Systems (CPS) allow for acting on the extracted intents from RQ1? Nurses at assisted living centers provide a helping hand literally and figuratively. Smart devices must serve this same function across various tasks (e.g., medication management).

## Significance of the Study

Human Activity Recognition (HAR) can improve elderly and special needs care by efficiently scaling out the visual coverage of medical facilities. Today, it is challenging to study HAR solutions within private residences. These issues stem from the system needing to record privacy-sensitive situations, such as bathing or intimacy, and then permit research students to examine the footage. Further complicating matters, the researchers must overcome the logistical challenges from finding representative samples, proving result reproducibility, and the economic overhead of multiple monitoring stations. Instead, this study proposes a research process using a physics simulator, animated actors, and virtual homes. The novel approach enables researchers to assess their CV algorithms across a repeatable configuration corpus. For instance, elderly patients falling is one of the most significant and avoidable reasons they need medical attention. This approach permits simulating this scenario, with each limb having distinct tensile strength, flexibility, and weight. When researchers can generate representative test-cases economically, it unlocks the potential for faster product iterations and expands the body of knowledge quickly.

Cyber-Physical Systems (CPS) serve as a bridge between digital algorithms and the real world. These technologies need patterns and methodologies that react to intents discovered through HAR. Today, the fractured ecosystem spans multiple vendors and is cumbersome to assess holistic solutions. This research project aims to reduce this complexity with specific virtual health and safety devices compatible with the simulator. Future researchers can leverage these tools and services to introduce noise (e.g., camera distortion) into the virtual world. Further lowering the barrier to entry for study HAR within personal residences opens the door to future innovations, not yet considered!

## Definition of Key Terms

### Actor

An actor is a 3D model that performs movement sequences inside the simulation environment. This study implements the construct using Filmbox (FBX) meshes and ROS character definitions.

### Action space

The CV algorithm observes an actor and predicts their intent (e.g., eating soup). Next, the intent maps into a taxonomy of known behaviors (called the action space).

### Artificial Intelligence/Machine Learning

Traditional software algorithms follow a model of . In contrast, AI/ML algorithms infer . This distinction permits the system to react to novel scenarios.

### Computer Vision (CV)

CV algorithms extract meaning from visual data, such as objects and labels from video streams.

### Convolutional Neural Network (CNN)

A family of AI/ML algorithms that perform computer vision.

### Cyber-Physical Systems (CPS)

A connected device communicates with a digital controller and responds through physical devices.

### Deep Neural Network (DNN)

A family of AI/ML algorithms maps a non-parametric function to a parametric function using a weighted network (Brown & White, 2017).

### Gazebo

An open-source 3D simulation framework for assessing actors and robots within a highly accurate physics environment (Bipin, 2018).

### Human Activity Recognition (HAR)

HAR is the process of mapping specific human behaviors to a known label (Gorgulu & Tasdelen, 2020).

### Internet of Things (IoT) device

### Motion capture (MoCap) modeling

### Recurrent Neural Network (RNN)

### Red Green Blue and Depth (RGB+D) format

### Robot Operating System (ROS)

### Wearable sensor

### World

## Summary