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 continually 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

The problem to be addressed in this study is mechanisms for improving elderly and special needs care through virtual assistance. Senior citizens live longer than ever before and want to defer moving into nursing homes until later in life. Similarly, special needs patients need the capabilities that supplement mental or physical deficiency. On the one hand, nurses can provide 24-hour supervision. This assistance could mean the difference between life and death during a fall and even routine tasks others take for granted. 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.

Medical facilities and family members address these situations through human capital investments, such as traveling nurses or guardian oversight (Westergren et al., 2021). It is prohibitively expensive for many patients to have dedicated staff. Instead, patients must compete for shared resources, increasing health care costs for some and degrading the quality for others. This divide disproportionally impacts lower-wage workers and minorities across the global population. This circumstance creates a need for technology and innovation to bridge the quality gap between remaining in residence or transitioning to a medical center. Capabilities to deliver this experience are possible by combining Artificial Intelligence/Machine Learning (AI/ML), Computer Vision (CV), and Cyber-Physical Systems (CPS).

Personal privacy, logistical complexity, ethical considerations, and economic overhead impede researching these solutions. Patients are hesitant to let academics record and persist their behaviors through real-time video monitoring technology. Furthermore, results are difficult to reproduce or extend.

## Purpose of the Study

The purpose of this constructive design research study is to create a research method for studying human behavior in privacy-sensitive contexts. It aims to deliver this capability utilizing humanoid constructs within a realistic physics simulation process. This approach could enable future researchers to assess their CV algorithms’ performance across diverse subjects rapidly.

Unlike humans, virtual actors are devoid of privacy and safety concerns while also duplicating economically. Furthermore, researchers can make their results reproducible externally through standard software change management procedures. These properties are directly applicable to lowering the barrier for investigating scenarios such as elderly and special needs care.

Next, the study demonstrates an implementation of critical aspects of the methodology. This step involves loading open-source Motion Capture (MoCap) animation sequences into the virtual world. Next, cameras will monitor the actor’s behaviors and predict its intent. Those intents will drive CPS virtual devices to mitigate any undesirable states, such as the patient has fallen or must perform a routine task (e.g., take medication). Lastly, the assessment process varies the actor’s configuration (e.g., weight, height, and flexibility). This feature set is essential to validate the AI/ML and CPS solutions generalizability.

Third, the study will identify mechanisms for scaling the simulation process over more extensive data sets. This assessment would consider scenarios such as increasing the action space, public cloud processing, and varying noise levels. The objective of this thought process is to determine future research areas and the next steps for commercialization. It is beyond the scope of this research project to deliver a production-grade simulation process, as the critical components are the research approach and demonstration of its application.

## Introduction to Theoretical Framework

Design-science is a research methodology that creates purposeful artifacts and applies them to study a phenomenon (Hevner, March, Park, & Ram, 2004). Both academic and business communities employ this method as a standard approach to Information Technology and Communication (IT&C) problems (Peffers et al., 2007; Bryar & Carr, 2021). It comes with well-defined guidelines (see Table 1) to implement a three-phased procedure. 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 study uses these guidelines and conceptual steps to identify a research-worthy topic and an actionable aspect. Next, it defines an abstract approach and implements a concrete proof-of-concept, the simulation process, to assess patient monitoring (via CV) and remediation (via CPS) technologies. Third, the artifacts expand the body of knowledge through the research questions. See Chapter 3: Research Method for more information.

## Research Questions

The foundational problem is elderly care and special needs patients need mechanisms that improve medical care within their homes. Medical facilities met this requirement through human capital investments (e.g., staff augmentation), which is expensive and difficult to scale. Instead, this study proposes that CV, AI/ML, and CPS technologies have the potential to supplement this need. This proposition raises questions regarding a potential solution’s ability to detect and respond to patient behaviors.

### RQ1

To what extent can CV-based systems extract the subject’s *intent* from dynamic and noisy video streams? Patients freely roam within their residence and modify the environment (e.g., move furniture and change lighting). As these factors influence the cameras’ perspective, does it inhibit the system from making accurate predictions, or do AI/ML algorithms account for these deviations?

### RQ2

To what extent can CPS-based systems react and mitigate the subject’s perceived behavior from RQ1? Medical staff at assisted living centers provide a helping hand literally and figuratively. Digital systems need to offer similar capabilities through IoT devices, which raises questions on the industry’s maturity level and commercial viability.

## 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

### Action space

Action spaces are the discrete or continuous responses that a subject can perform (Lapan, 2018). In the 1983 classic game, Mario Bros. (Wikipedia, n.d.), the main character has an action space of moving left or right, jumping, and throwing a fireball.

### Artificial Intelligence/Machine Learning (AI/ML)

AI/ML algorithms can generalize business rules from data for predicting outcomes to novel examples (Buchanan, 2005). This structure differs from traditional algorithms, which produce outputs (data) versus business rules.

### Computer Vision (CV)

Computer (or machine) vision is a set of capabilities that extract information from 2D and 3D images (Hornberg, 2017).

### Convolutional Neural Network (CNN)

CNN algorithms are specialized DNN architectures that predict outcomes from image sources (Nguyen, Huynh, Tran, & Ngo, 2019).

### Cyber-Physical Systems (CPS)

Cyber-Physical Systems are network-programmable devices that respond to digital messages through embedded capabilities (Aguida, Ouchani, & Benmalek, 2020). It is a subset of an Internet of Things (IoT) device.

### Deep Neural Network (DNN)

Neural Networks map a non-parametric function to a parametric function using a forward feeding weighted graph (Brown & White, 2017). Sophisticated situations require an architecture with multiple internal mapping-to-mapping constructs (called hidden layers). These *deep* networks gain specialization from the hidden layers, such as predicting edges into figures into body parts (Fridman, 2017).

### Gazebo

The Gazebo framework is an open-source simulation process for assessing actors’ and robots’ performance through a physic engine (Bipin, 2018). This application is a standard utility for many simulation workloads.

### Human Activity Recognition (HAR)

HAR is the process of mapping specific human behaviors to known labels (Gorgulu & Tasdelen, 2020). It makes these predictions using CNN and RNN algorithms to evaluate changes in image data over time.

### Internet of Things (IoT) device

The Internet of Things (IoT) attempts to widen the interconnectivity of computers by interconnecting objects (Commission of the European Communities, 2009). These objects expose sensors connected to web services to provide personalized data feeds.

### Motion capture (MoCap)

Motion capture is a process that digitizes structural movement for film and television production (Gan, Li, Wang, & Zhang, 2020). Practitioners collect this information using tracking points on actors that perform specific behaviors (e.g., walking).

### Recurrent Neural Network (RNN)

Recurrent Neural Networks specialize in making predictions on sequential data sets like natural language processing and time series (Boorugu & Ramesh, 2020).

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

Microsoft Kinect and similar cameras encode image data with color and depth channels (Litomisky, 2012). This additional depth channel simplifies actor movement tracking within 3D space.

### Robot Operating System (ROS)

ROS is a meta operating system that standardizes communication between heterogeneous agents (Bipin, 2018). It exposes core shared services like agent state management, message passing constructs, connectivity to simulated processes, and interfaces for physical hardware.

### Unreal Engine

UE is an extensive content creation suite with numerous agent simulation capabilities (Unreal Engine, 2022). It exposes features for realistic physics and dynamic world modification.

### World

A world refers to the virtual simulation process that contains the various actors and inanimate objects within the test (Bipin, 2018). Some game theory articles also refer to this construct as a “level.”

## Summary

This chapter introduced the central problem statement that researchers need to study privacy-sensitive contexts like in-home monitoring of elderly and special needs patients. Those researchers encounter challenges spanning personal privacy, logistical, sufficient and diverse representation, and costs, among other entry barriers. After mitigating these issues, the research results are difficult and expensive to reproduce.

The study proposes reducing these constraints through a simulation process with

and virtual research method for observing virtual actors within a simulation process. Researchers can configure the actors to perform specialized animation sequences under different physical parameters (e.g., height and weight). Those values control a physics engine that makes real-time customizations to the scene. Next, the study implements this research method using existing open-source software and MoCAP repositories. Then, it positions virtual cameras within the environment to perform HAR tasks.