Week 8: An unobtrusive elderly care system

Nate Bachmeier

TIM-7245: Directed Constructive Research

August 15, 2021

Northcentral University

Table of Contents

[An unobtrusive elderly care system 5](#_Toc79692903)

[Problem Statement 5](#_Toc79692904)

[Purpose Statement 6](#_Toc79692905)

[Research Questions 7](#_Toc79692906)

[Hypotheses 8](#_Toc79692907)

[Definition of Key Terms 9](#_Toc79692908)

[Theoretical Framework 9](#_Toc79692909)

[Fundamental Approach 10](#_Toc79692910)

[Research Artifact Design 11](#_Toc79692911)

[Core Architectural Design 11](#_Toc79692912)

[Extracting Intents 11](#_Toc79692913)

[Data Collection and Monitoring 11](#_Toc79692914)

[Feedback Mechanism 12](#_Toc79692915)

[Tenants, Controversies, and Ethical Challenges 12](#_Toc79692916)

[Literature Review 13](#_Toc79692917)

[Who is the customer 13](#_Toc79692918)

[Challenges with current solutions 13](#_Toc79692919)

[Implementing unobtrusive systems 14](#_Toc79692920)

[Scaling the classification process 16](#_Toc79692921)

[Integrating AI/ML into the ecosystem 16](#_Toc79692922)

[Including Cyber-Physical Systems 17](#_Toc79692923)

[Summary 19](#_Toc79692924)

[Customer Observations 19](#_Toc79692925)

[Unobtrusive Methods 19](#_Toc79692926)

[Integrating Cyber-Physical Systems 19](#_Toc79692927)

[Research Methods 20](#_Toc79692928)

[Artifacts 20](#_Toc79692929)

[Contributions 21](#_Toc79692930)

[System Architecture 22](#_Toc79692931)

[User Roles 22](#_Toc79692932)

[System Reliability 23](#_Toc79692933)

[Measurements and Evaluation 23](#_Toc79692934)

[Data Collection Process 23](#_Toc79692935)

[Evaluation Process 24](#_Toc79692936)

[Benchmarking 25](#_Toc79692937)

[Summary 25](#_Toc79692938)

[References 27](#_Toc79692939)

[Annotated Bibliography 30](#_Toc79692940)

[Human Activity Recognition 30](#_Toc79692941)

[Integrating IoT Systems 31](#_Toc79692942)

[Enhancing Security 31](#_Toc79692943)

[Healthcare and Cloud 32](#_Toc79692944)

# An unobtrusive elderly care system

This constructive research project seeks to raise the elderly care quality bar while minimizing costs. Today, many researchers are approaching this problem through wearable technologies. However, this approach is not practical for several reasons, such as requiring the person to carry the device at all times. Instead, Human Activity Recognition (HAR) through computer vision and machine learning can provide equal or better coverage without being obtrusive to many scenarios. After observing the patient’s behavior, the solution can orchestrate Cyber-Physical Systems (CPS) to provide further assistance.

## Problem Statement

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. Alternatively, researchers are exploring wearable IoT devices (Tun et al., 2021). These sensors provide mechanisms for requesting assistance and receiving continuous monitoring. Nevertheless, there are many limitations to wearable solutions. Most notably, the person must *remember* to wear them, which raises challenges for early-onset memory loss patients.

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, it enables the patient to 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). Medical facilities can address these challenges through real-time video monitoring services that analyze patients’ actions and recommend care. For instance, patients with memory impairment might forget to empty the dishwasher, take medication, or bathe regularly. These scenarios are challenging to address through wearable devices. However, an in-home camera system becomes a watchful eye that spots those missing actions through computer vision. After detecting an issue, the system alerts the person using Text-to-Speech (TTS) services (e.g., Amazon Alexa) and other Cyber-Physical Systems (CPS).

## Purpose Statement

This constructive design research project defines and implements an Elderly Care Smarthome Operating System (ECSOS). The ECSOS will provide central core services for bringing world-class assisting living care into a resident’s home, such as identity management, patient action tracking, consistent Cyber-Physical control plane, and privacy functions. 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.

Building these capabilities requires tooling spanning networking, sensors, embedded systems, and real-time video processing (Elloumi et al., 2020; Das et al., 2019). This research will leverage industry-standard tooling (e.g., JavaScript, Apache Spark, and Tensorflow). In addition, specific aspects necessitate custom code that enhances existing open-source software (e.g., IoT control interfaces and Python libraries). Together, these different technologies culminate into an elegant solution that monitors, predicts, and responds in real-time to patient needs.

Next, a case study will assess the solution’s effectiveness against alternative approaches (e.g., wearables). This phase requires installing WiFi cameras and collecting example footage. Finally, the study participants will give qualitative prediction accuracy feedback (e.g., 1-5 star scores). Their responses combine with various statistical metrics (e.g., number of predictions) to produce a holistic system assessment.

## Research Questions

Researchers are innovating across health care using Internet of Things (IoT) devices. Their efforts predominately focus on wearable technologies that attach sensors to the patient (Tun et al., 2021). Wearable technologies face significant competition because these solutions have a low barrier to entry, economical pricing, and are mass-producible. However, these products lack elegance due to restricting movement and necessitating the patient always to carry the device. Additionally, the saturated market causes each iteration to produce less incremental value.

In contrast, high-quality research must be challenging, elegant, and move the needle (Zeller, 2014). Meeting these requirements necessitates a different approach, such as utilizing cameras and real-time video processing to deliver a superior experience. However, video-centric systems encounter more complexity in several aspects. For instance, patients can freely move around their residence and change its configuration (e.g., move furniture or turn off a light). Addressing the noise within these dynamic environments is challenging and creates multiple research questions.

**R1** – How are researchers minimizing noise in their video streams? An efficient process must exist to analyze short videos and extract the subject’s *intent*. This mechanism must reliably handle noisy data (e.g., out-of-focus images) and variable input (e.g., distance to the camera).

**R2** – How can the extracted intents best *interface* with Cyber-Physical Systems (CPS)? Nurses at assisted living centers provide a helping hand literally and figuratively. Smart devices must serve this same function across a range of tasks (e.g., medication management).

**R3** – How can those interfaces ensure patient *confidentiality*? Patients will only use a continuous video recording solution if they trust its security and privacy controls. There must be explicit and deliberate decisions regarding how information is stored and replicated.

**R4** – How can central administrative teams most efficiently *scale* across global and domestic sales territories? Healthcare workers can remotely deliver world-class services because the homes contain CPS systems for routine tasks (e.g., monitoring patient falling). Competitive businesses can leverage this capability to decrease costs, increase profit margins, and maintain quality standards.

## Hypotheses

There are existing mechanisms to address each system requirement. However, those capabilities exist as isolated components. An Elderly Care Smarthome Operating System (ECSOS) can create a consistent control plane that brings these services together. The ideal driver of such a solution is real-time video processing at the edge. This approach promotes security and privacy by minimizing external data transfers. Additionally, small appliances built on the Raspberry PI platform have sufficient resources to drive computer vision models. Presumably, combining these different technologies will improve patient care, deliver more value, and lower costs.

## Definition of Key Terms

This constructive research proposal uses several industry-standard terms. Table 1 contains a brief list for reference.

Table 1: Key Terms

|  |  |
| --- | --- |
| Term | Definition |
| **Computer Vision** | Algorithm(s) that rationalize image contents |
| **Convolutional Neural Network** | Algorithm(s) for object detection within images |
| **Cyber-Physical Systems (CPS)** | Mechanisms for bridging computer algorithms into the physical world |
| **Elderly Care Smarthome Operating System (ECSOS)** | A consistent control plane, automation, and devices that manage a residence |
| **Ensemble Algorithm** | Algorithm(s) that composite multiple child algorithms into a collective output |
| **Function as a Service** | A stateless compute model that optimizes for short-lived tasks |
| **Human Activity Recognition** | The ability to classify body movements as actions |
| **Machine Learning Model** | A weighted decision tree that predicts an outcome |
| **Neural Network** | A mechanism for approximating a non-linear function from high-dimensional data |
| **No(t only) SQL** | Data structure management services that support specific access patterns |
| **Public and Private Clouds** | A collection of compute and storage resources with programmable interfaces |
| **Recurrent Neural Network** | Algorithm(s) for detecting patterns in data sequences |
| **RGB+D** | Red, Green, Blue, and Depth encoded image |
| **Wearable Technology** | Connected devices and sensors that collect data by attaching to the person |

# Theoretical Framework

There are four approaches to studying a business use-case or phenomena (see Table 2). Constructive design is one of the most common research methods for information systems and technology (Silvestrini & Sammito, 2012). These studies identify a problem, build artifacts, and communicate the implementation’s unique value (Hevner et al., 2004). In addition, many researchers follow this process to build proof-of-concept and execute case studies. This methodology is appropriate for creating an Elderly Care Smarthome Operating System (ECSOS) solution. After building the system, it can support a targeted case study that measures its abilities to deliver value.

Table 2: Example Research Strategies for Classifying Movement in Video

|  |  |  |
| --- | --- | --- |
| Approach | Description | Study Example |
| Quantitative | Studies the magnitude of a phenomena | Measure the resources necessary to classify movement with embedded systems |
| Qualitative | Explores a concept without a numerical basis | Exploration of reasons movement classification fails |
| Mixed-Method | Combines exploration and studying the magnitude of these issues | What preparation steps reduce the costs movement classification |
| Constructive | Produce artifacts to study a scenario | Create an algorithm for classifying movements |

## Fundamental Approach

Constructive research practitioners gravitate toward either Design Science Research (DSR) or the Constructive Research Approach (CRA). One of the critical differences between them is that DSR relies more heavily on existing theories, versus CRA does not explicitly require a base theory (Piirainen & Gonzalez, 2013). More recently, Iivari (2020) criticizes the debate stating that constructive research must first and foremost produce high-quality artifacts. She advocates for “less theory, but better design theory (pg. 504),” especially within rapidly evolving industries like Information Technology and Communication. Zeller (2014) would agree with this position, adding success criteria that the artifacts are “challenging, elegant and useful.” This research project aligns with these requirements through its focus on connecting artifacts with customer needs and challenges. For instance, it would be easier to “create another smartwatch” that does not adequately serve the user base.

## Research Artifact Design

This research project has four core components which collectively form a proof-of-concept implementation and a mechanism to measure results. According to the literature review, these systems have the potential to output high-quality constructive research.

### Core Architectural Design

First, the team installs WiFi-enabled Eufycam 2C cameras to collect short recordings. These cameras use motion-sensing to trigger short Audio/Video (A/V) recordings (fifteen to sixty seconds). After the filming completes, its controller (Eufy Homebase) automatically uploads the file to Network Attached Storage (NAS). The file creation event triggers an analysis workflow that extracts and publishes metadata to message buses. Developers can author extensions using Function-as-a-Service (FaaS) constructs that subscribe to the notifications.

### Extracting Intents

Second, a machine learning algorithm will classify and annotate the video’s contents. There are several potential implementations (e.g., Open Pose versus Toyota’s approach). The performance and resource requirements between these strategies require analysis. Ideally, the AI model can run in an edge appliance versus uploading into a Public Cloud Service (PCS). However, this raises concerns that the device has sufficient computing capabilities (e.g., parallel processing dozens of cameras). If analysis occurs within the cloud, then it introduces security and privacy concerns.

### Data Collection and Monitoring

Third, the ECSOS solution routes the metadata into monitoring and response extensions. These extensions include central services (e.g., identity and state management) and auditing capabilities (e.g., inputs, predictions, and recommendations). One crucial extension is the central audit logs. These tables are queryable within a NoSQL time-series database (e.g., Influx).

### Feedback Mechanism

Fourth, users can provide feedback on model predictions through a mobile app. The feedback contains qualitative tags and centrally aggregates for offline analysis. The research team will use this database to assess the system’s accuracy and identify potential quality gaps. For instance, the responses could indicate that specific intent predictions are unreliable (e.g., use fork versus pen). Those issues require prioritization and potentially additional training data.

## Tenants, Controversies, and Ethical Challenges

The system’s primary purpose is to increase the patients’ quality of life by remaining within their residency longer. Therefore, this mission statement obliges the solution to detect human activity and respond reliably. Also, patients will only use a continuous video recording solution if they trust its security and privacy controls. There must be explicit and deliberate decisions regarding how information is stored or transferred.

There are four major threats to research validity internal, external, statistical conclusions, and construct validity (see Table 2). According to Parker, “it is widely accepted truism that all published research to some extent is flawed. Because the research enterprise is fraught with many pitfalls, researchers must become well-versed in recognizing, and when possible avoiding design shortcomings (Parker, 1993, p. 1).” The research must explicitly implement controls that minimize these concerns and ensure the results are reproducible. For instance, the video recordings need to include diverse subjects and automation for external quality auditing.

Table 3: Threat Sources

|  |  |
| --- | --- |
| Source | Description |
| Internal Threat | Contamination by the research team |
| External Threat | Contamination outside of the study’s controls |
| Statistical Conclusion Validity | Results are arbitrary or non-reproducible |
| Construct Validity | Controls are not enforceable or consistent |

# Literature Review

## Who is the customer

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 adequate services availability. 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 an assisted living home.

## Challenges with current solutions

Inversely, the explosive growth across IoT, Cloud, Big Data, and Mobile (ICBM) continuously decreases costs and enables new scenarios. These technologies will 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 available and stifling creativity.

Technology within special needs and elderly care settings has unique challenges and requirements (Ferati, Kurti, Vogel, & Raufi, 2016). For example, these solutions often support safety functions and require uninterrupted monitoring. Wearable IoT has inherent risks that the patient will disable or forget the device. This situation is particularly profound in patients with memory impairments like dementia (Wilson, 2017). Additionally, the obtrusive nature of wearable technologies makes them impractical for extended duration scenarios (Razzaq et al., 2020; Singla et al., 2010). Alternatively, specific vendors utilize voice-enabled Personal Digital Assistants (PDA) (e.g., Amazon Alexa). These products effectively set reminders and record activities (Tan et al., 2020). However, several scenarios cannot exploit vocal interactions, such as 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 timeslices. Instead, patients could receive continuous observation at lower costs using Computer Vision. Artificial Intelligence and Machine Learning (AI/ML) models can observe patient behavior and react accordingly.

## Implementing unobtrusive systems

Amazon Go, a cashier-less store, proves the potential through sophisticated computer vision technologies that can even protect against shoplifting (Wankdhede et al., 218). However, training these computer vision models requires domain-specific video clips (Das et al., 2019; Razzaq et al., 2020). Most open-source datasets contain outdoor and sporting activity, which is generally different from indoor behavior. Researchers mitigate these issues by minimizing project scope (Yi & Feng, 2021; Chen et al., 2020). While this is acceptable during the proof-of-concept design, it will impact productization.

After creating the training video, perform Human Activity Recognition (HAR) through a two-step process (Razzaq et al., 2020; Chen et al., 2020; Yi & Feng, 2021). First, a process extracts the subject’s skeletal position from a given frame. This phase begins with decoding the information into RGB+D (Red Green Blue and Depth) matrices. Next, the researchers use Convolutional Neural Networks (CNN) to identify the body parts within the frame. If the room has multiple cameras, then a normalization process must build a consistent 3-D representation. The normalization needs to also account for noise and scaling challenges. These issues occur because the patient can freely move around the room. Microsoft’s Kinect Sensor supports automating this step to a certain extent. Alternatively, Carnegie Mellon University’s Open Pose library can approximate the 3-D skeletal structure from 2-D images.

Figure 1: Extract Skeletal Process

Second, a process needs to determine the skeletal changes by comparing the input 3-D skeleton against the previous frame’s position. After determining the delta movement updates, a recurrent sequence algorithm predicts the activity. The implementation is different between each publication (e.g., Dynamic Time Warping and Long-Term Short-Term Memory LTSM). However, the general strategy of using Recurrent Neural Networks (RNN) remains the same. After evaluating a sequence of skeletal movements, the algorithm will classify the action.

Figure 2: Movement Classification

## Scaling the classification process

It would be challenging to build one movement classification system that supports the entire universe of human activities. Instead, production systems blend hundreds of models into ensemble algorithms (Bell, Koren, & Volinsky, 2009). For example, predicting that the subject is picking up a fork, spoon, or cup is similar. The classifier can determine which subactivity compositing predictions the subject lifts an object and a standard object classifier.

Though, in practice, there can be issues programmatically choosing the correct composite models. One solution involves executing dozens of binary classifiers, then filtering the outputs using a voting scheme (Delgado, 2021). For example, three binary classifiers might exist to predict the subject is lifting an object. Then under a simple majority vote, the meta classifier equals the decision with at least two votes. System designers might require multiple voting schemes across the solution depending on the criticality of being right. For instance, if three classifiers are predicting the patient has fallen, one vote might be sufficient.

## Integrating AI/ML into the ecosystem

The most powerful artificial intelligence applications use machines to enhance human capabilities rather than replace them (Heer, 2019; Boire, 2017). For instance, a person can write a more profound business case than a machine; however, the same machine will have fewer misspellings and grammatical errors. This dichotomy exists because humans specialize in contextualizing thought versus using patterns to make predictions (Schleer et al., 2019). Artificial intelligence is a tool that can automate mechanical tasks, pattern match data, and enhance human capabilities. Organizations can use these means to improve efficiency and reduce wastefulness. These innovations deprecate the need for specific skill sets and lower the entry barrier into other expert systems. While this causes an initial decrease number of jobs, entirely new industries follow shortly afterward. When a society can replace low-paying jobs with multiple high-paying industries, this promotion justifies the short-term pain.

For instance, a recent demonstration shows that AI/ML can map facial expressions to pain levels with 93% accuracy (Nugroho et al., 2018). These capabilities could enable every patient access to a well-trained and vigilant private nurse. Other publications chose to forward the observations to cloud services for remote management features (Elloumi et al., 2020; Chen et al., 2020). After centralizing patient telemetry, the nursing staff can more efficiently prioritize their time and resources.

## Including Cyber-Physical Systems

The Internet of Things (IoT) represents the next evolutionary step in communication and system connectivity. Naïve outsiders see this industry as a series of gimmicks, Apple watches, and Smart toasters. More importantly, those statements are factual and create the missing bridge between cyber and physical systems (CPS). This capability comes from sensor and input networks that emit telemetry into ubiquitous cloud computing and machine learning platforms. Big Data and artificial intelligence solutions can control manufacturing and safety systems using physical motors and actuators. As information and decision processes transact across this bridge, organizations can execute expert workflows autonomously and prevent costly failures.

Health care solutions can leverage these bridges to offload specific nursing tasks to autonomous devices. For example, the computer vision processes might detect the patient is experiencing mild pain and dispense an aspirin (Amin, Salahuddin, & Bouras, 2020). Before dispensing medication, the CPS can optionally request authorization from the central nurse’s station. This design flexibility enables administrators to set guardrails and ensure consistency.

While many health care CPS scenarios are easy to describe, they often span complex workflows (Abdulameer et al., 2015). For instance, the aspirin dispenser involves computer vision, edge processing, cloud computing, manual approval steps (optional), and finally, orchestrating a medication vending machine. These heterogeneous components span distinct vendors, protocols, and technology stacks, making building secure and reliable service integrations difficult (Riesener et al., 2021). Researchers must be cognizant of these implementation challenges and minimize additional complexity.

One approach is through multi-layered micro-service architectures, Schema-on-Read semantics, and formalizing data consistency procedures (Aguida, Ouchani, & Benmalek, 2020). However, distributed systems are the most complicated computing environments because of their parallel and asynchronous nature. Many implementations also make false assumptions regarding the network’s reliability, security, homogeneousness, latency, bandwidth, and transport costs (van Steen & Tanenbaum, 2019, p. 986). When systems introduce one of those fallacies into the design, it produces subtle defects under production loads. Businesses combat these risks through high-availability architectural patterns that promote self-healing (Yang, Min, Yang, & Li, 2014). These strategies follow reactive (e.g., heart-beating) combinations and proactive solutions (e.g., rejuvenation tactics).

## Summary

The literature review encompasses three core ideas—first, an examination of the customer and current solutions. Next, an assessment of state-of-the-art unobtrusive methods concludes that Human Activity Recognition is the dominant strategy. Lastly is a review of digital-to-physical methods within healthcare scenarios. Collectively these domains confirm that the ECSOS solution unobtrusively solves the right problem with reactive capabilities within the physical world.

### Customer Observations

Several global communities could benefit from assisted living facility-like capabilities within their residence. This cohort includes senior citizens, people with disabilities, and young children. Today, researchers predominately meet these requirements through wearable technologies. The wearable technologies market is commoditized and overly saturated, which limits innovation. Further, this approach is obtrusive and impractical for many indoor scenarios.

### Unobtrusive Methods

Implementing unobtrusive systems is significantly more complex, requiring a collection of Computer Vision and Human Activity Recognition (HAR) algorithms. Researchers typically build a process that tracks skeletal structures in 3-D space using Convolutional Neural Network (CNN) algorithms. Afterward, Recurrent Neural Networks (RNN) monitor changes to predict the patient’s behavior. Maintaining massive monolithic architectures is challenging, and this causes practitioners to employ ensemble methods.

### Integrating Cyber-Physical Systems

Requiring registered nurses to perform every task is prohibitively expensive. Instead, automation must exist for, e.g., dispense medication. Those systems require a communication bridge that enables cloud services to manipulate the physical world. Several researchers explore these ideas through targeted case studies and mock stages (e.g., miniature replica houses). These concept studies enumerate numerous successes and challenges with integrating health care into a private residence. Generally speaking, the research suggests that existing technology is adequate for implementing these systems. However, designers must cobble components from multiple vendors that utilize different technology stacks. Without a consistent control plane and data models, the barrier to entry is significant for engineering teams.

# Research Methodology

Before starting any significant undertaking, there needs to be a formal project plan that scopes the intent. Additionally, the plan must define mechanisms to measure the success and impact of those efforts. Without those prerequisites, it can be challenging or impossible to prove the resources efficiently produce the project’s results.

## Artifacts

This research project has three core components which collectively form a proof-of-concept implementation and mechanism to measure results. First, the team installs WiFi-enabled Eufycam 2C cameras to collect short recordings. These cameras use motion-sensing to trigger short Audio/Video (A/V) recordings (fifteen to sixty seconds). After the filming completes, its controller (Eufy Homebase) automatically uploads the file to Network Attached Storage (NAS). The file creation event triggers an analysis workflow that extracts and publishes metadata to message buses. Developers can author extensions using Function-as-a-Service (FaaS) constructs that subscribe to the notifications.

Second, a machine learning algorithm will classify and annotate the Human Activity Recognition (HAR) metadata. There are several potential implementations (e.g., Open Pose versus Das et al.’s approach). The performance and resource requirements between these strategies must exist. Ideally, the model can run in an edge appliance versus uploading into a Public Cloud Service (PCS). However, this raises concerns that the device has sufficient computing capabilities (e.g., parallel processing dozens of cameras). If analysis occurs within the cloud, then it introduces security and privacy concerns.

Third, the ECSOS solution routes the metadata into monitoring and response extensions. These extensions include central services (e.g., identity and state management) and auditing capabilities (e.g., inputs, predictions, and recommendations). One crucial extension is the central audit logs. These tables are queryable within a NoSQL time-series database (e.g., Influx).

## Contributions

The core contribution to the body of knowledge is the case study using the proof-of-concept design. Existing publications review each component within ECSOS under distinctly different use cases (e.g., sports injuries). Das et al. (2019) explain that those resources are not directly reusable, and implementations must use domain-specific labeled content. This design requirement necessitates compositing a new solution from custom and open-source software.

Second, the research produces a purpose-built machine learning algorithm for elderly care action recognition. This deliverable also includes quantitative metrics that describe the algorithm’s resource utilization and F-measure accuracy. Data scientists use F-measurements as a “way of combining the precision and recall of the model, and [defines] the harmonic mean of the model’s precision and recall (Wood, n.d.).” Researchers can make trade-offs in their solution to optimize this value for their specific scenario. For instance, a critical health management system might enforce higher penalties on false negatives than over positives.

## System Architecture

Elderly Care SOS requires cameras, network storage, and a custom-built appliance (see Figure 1). Optionally patients can extend the system with various CPS device integrations (e.g., remote smoke detector). The appliance must have enough computing and storage resources to perform model predictions, persist state, and execute several micro-services. Periodically, the on-premise system needs to synchronize with an external cloud component. These synchronization operations include sending status reports, downloading updates, and issuing assistance requests.

Figure 3: Abstract Design

Diagram

Description automatically generated

## User Roles

Aside from patients, there are three additional user roles: nurses and healthcare providers, family members, and administrators. These users can use a mobile app or web portal to access the relevant data. All operations from either UI (User Interface) require Authentication, Authorization, and Auditing (AAA). When systems mandate AAA enforcement, it prevents negligence or malicious actions while increasing transparency. It is also critical that the patient maintains control of their privacy. For example, they might want to share a weekly aggregate health report with family members, not verbose details.

## System Reliability

The architecture’s components communicate over WiFi, Zigbee, and Bluetooth protocols. These messages are likely to encounter transfer failures due to radio interference or devices being offline. There must be support within the message buses to cache and reattempt any message delivery failures using exponential backoff policies. Otherwise, the state management’s perspective can become distorted. Time-sensitive messages (e.g., the subject has fallen) require a primary and secondary communication channel, such as phone line or mobile phone pairing.

The appliance must locally run several services that handle core scenarios like identity and message routing. Developers can also load custom extensions that subscribe to event notifications. Those various subsystems require isolation and controls to limit the blast radius of a specific failure. An industry-standard approach would be to use micro-service designs and container orchestration technologies (e.g., Kubernetes)(Wen et al., 2020). These products can manage fail-over replicas and promptly restart crashed instances.

# Measurements and Evaluation

## Data Collection Process

Most information enters the system through the WiFi cameras. Ideally, those cameras are accessible only through a dedicated Virtual Local Area Network (VLAN). This recommendation protects the unencrypted Real-Time Streaming Protocol (RTSP) from eavesdropping and tampering attacks. After the video clip is available, ECSOS must process it through several machine learning models (e.g., facial recognition, object detection, and action recognition). These metadata annotations persist into a time-series database. Lastly, populating the database requires the patient(s) to behave normally and let the system collect the video recordings.

Initially, the system will not have any training data and cannot make recommendations. However, researchers can accelerate data labeling with online products like Amazon SageMaker Ground Truth (Amazon, 2021). This service offers clustering capabilities to group related artifacts and streamlines manual tasks. Alternatively, users can crowd-source labeling jobs through Amazon Mechanical Turk.

## Evaluation Process

There must be feedback loops that confirm that the predictions and recommendations are accurate. Without this capability, it would be challenging to discover issues and prioritize machine learning model changes. One potential solution is to collect these responses through a patient mobile app. App users can enumerate previous recordings and see the associated metadata. Those filmings are subject to a retention policy that automatically deletes old content. If they disagree with the predictions, they can make corrections inline. After making the manual update, the user can help improve the experience by submitting the footage to the ECSOS Cloud.

The ECSOS Cloud maintains a history of all incorrect predictions. Data scientists can review those responses, look for patterns (e.g., mixing up two actions), and make the necessary modifications. There must be some mechanism to include user feedback and avoid introducing biases into the model (García-Pérez, 2012). For example, an individual user could post thousands of feedback comments containing inaccurate data change requests. This situation could regress other user’s experiences. Similar biases can enter the system due to insufficient test subject’s racial diversity (e.g., only validating white men).

## Benchmarking

Numerous Human Activity Recognition (HAR) benchmarks exist with varying frame rates, actions, actors, backgrounds, resolutions, and domains (Singh & Vishwakarma, 2018). However, most benchmarks also focus on high-intensity outdoor sports footage (Das et al., 2019). Since those behaviors are very different from low-intensity indoor movements, they are not directly usable. Instead, several publications choose to define movement taxonomies and curated lists of expected behaviors. Afterward, the project’s quality is proportional to its ability to cover those actions. Additionally, the solution should be extensible and support more actions over time (e.g., future versions).

# Summary

Senior citizens and other populations are living longer and want to defer transitioning into assisted living facilities. Delaying the move can save nearly ninety thousand dollars annually and additional value (e.g., personalized comfort). However, those savings typically require care degradation. Researchers and designers are mitigating these issues through wearable technologies (e.g., smartwatches). These solutions are impractical for extended periods due to restricting movement, requiring carrying the device, remembering to charge it, among other limitations. Instead, the state of the art solutions utilizes real-time video processing to provide an equal or better experience without being obtrusive.

Computer vision strategies typically employ a skeletal tracking process. This process consists must first convert the patient’s skeleton into 3D space then assess the changes. That operation requires a combination of CNN, RNN, and ensemble algorithms. Multiple challenges arise while implementing these pipelines. For instance, there are few open datasets for analyzing indoor low-activity behaviors. This limitation causes researchers to build everything from step one forward. Luckily, there are existing tools and services to simplify the development experience and design complex behaviors.

After determining the patient’s *intent*, the operating system must respond with an appropriate response. This reaction frequently requires manipulating the physical world from the virtual universe. Cyber-Physical Systems create those communication bridges, enabling data analytics to orchestrate physical devices. There are several well-established patterns for approaching these challenges, but the ecosystem is fractured (e.g., requires multiple vendors).

Lastly, unlocking these challenges could enable an extremely competitive service offering. Today, assisted living centers must employ numerous staff members and maintain the physical buildings. Instead, the ECSOS enables centralized nursing teams to scale across sales territories, reducing costs and increasing efficiencies. Additionally, since the patients remain within their residence, the business can leverage smaller facilities. Furthermore, the service is marketable to a broader audience. While traditional living centers only care for end-of-life patients, ECSOS applies to any widow or elderly parent. Children are often hesitant to “put their mom in a home” but gladly subscribe to the piece of mind that “someone is watching mom.”

# References

Aguida, M., Ouchani, S., & Benmalek, M. (2020). A review on cyber-physical systems. *Enabling Technologies: Infrastructure for Collaborative Enterprise* (pp. 275-278). Virtual: IEEE. doi:10.1109/WETICE49692.2020.00060

Amazon. (2021). *Use Aamzon SageMaker Grouth Truth to Label Data*. Retrieved from Amazon: https://docs.aws.amazon.com/sagemaker/latest/dg/sms.html

Amin, S., Salahuddin, T., & Bouras, A. (2020). Cyber-Physical Systems and Smart Homes in Healthcare: Current State and Challenges. *International Conference on Informatics, IoT, and Enabling Technologies* (pp. 302-309). Virtual: IEEE. doi:10.1109/ICIoT48696.2020.9089638

Bell, Koren, & Volinsky. (2009). *The BellKor solution to the Netflix Prize.* Retrieved from Netflix Prize: https://netflixprize.com/assets/GrandPrize2009\_BPC\_BellKor.pdf

Das, S., Dai, R., Koperski, M., Minciullo, L., Garattoni, L., Bremond, F., & Francesca, G. (2019). Toyota Smarthome: Real-World Activities of Daily Living. *International Conference on Computer Vision* (pp. 833-842). Seoul, Korea: IEEE. doi:10.1109/ICCV.2019.00092

Delgado, R. (2021). A semi-hard voting combiner scheme to ensemble multi-class probabilistic classifiers. *Applied Intelligence, 1*, 1-25. doi:10.1007/s10489-021-02447-7

Ferati, M., Kurti, A., Vogel, B., & Raufi, B. (2016). Augmenting requirements gathering for people with special needs using IoT. *International Workshop on Cooperative and Human Aspects of Software Engineering* (pp. 48-52). ACM. doi:10.1145/2897586.2897617

García-Pérez, M. A. (2012). Statistical conclusion validity. *Frontiers in Psychology, 3*. doi:10.3389/fpsyg.2012.00325

Koreshoff, T., Robertson, T., & Leong, T. (2013). Internet of Things: a review of literature and products. *CHI’13, November 25 - 29 2013, Adelaide, Australia*.

Morris, J. (2008). *Disability research and policy: current perspectives.* Lawrence Erlbaum Associates.

Muhsin, A., Munyogwa, M., Kibusi, S., & Seif, S. A. (2020). Poor knowledge on elderly care despite positive attitude among nursing students in Zanzibar Island: findings from a cross-sectional study. *BMC Nursing, 19*(1), 1-8. doi:10.1186/s12912-020-00488-w

Parker, R. (1993). Threats to the validity of research. *Rehabilitation Counseling Bulletin, 36*(3), 130-138. Retrieved from https://search-ebscohost-com.proxy1.ncu.edu/login.aspx?direct=true&db=eric&AN=EJ458938&site=eds-live

Piirainen, K., & Gonzalez, R. (2013). Constructive Synergy in Design Science Research: A Comparative Analysis of Design Science Research and the Constructive Research Approach. *Liiketaloudellinen Aikakauskirja, 3*(4), 206-234. Retrieved from https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib&db=bth&AN=95116694&site=eds-live

Silvestrini, R. P., & Sammito, G. (2012). Design of Experiments for Information Technology Systems. *Defense AT&L, 41*(5), 30-35. Retrieved from https://search-ebscohost-com.proxy1.ncu.edu/login.aspx?direct=true&db=bth&AN=80409129&site=eds-live

Singh, T., & Vishwakarma, D. (2018). *Human Activity Recognition in Video Benchmarks: A Survey.* Springer.

US Bureau of Labor Statistics. (2020, May). *Registered Nurses*. Retrieved from US Bureau of Labor Statistics: https://www.bls.gov/ooh/healthcare/registered-nurses.htm

Wilson, C. (2017). *Caring for people with dementia: a shared approach.* Los Angeles, California: SAGE.

Wood, T. (n.d.). *What is the F-score*? Retrieved from Deep AI: https://deepai.org/machine-learning-glossary-and-terms/f-score

Yang, M., Min, G., Yang, G., & Li, Z. (2014). Software rejuvenation in cluster computing systems with dependency between nodes. *Computing, 96*, 503–526. doi:10.1007/s00607-014-0385-x

# Annotated Bibliography

Examining the Northcentral University (NCU) Library with search terms such as elderly care, IoT, and video health monitoring uncovers several industry-wide trends.

## Human Activity Recognition

Amazon Go enables customers to purchase goods from physical stores without requiring cashiers (Amazon, 2021). Their solution uses Deep Neural Network (DNN) algorithms that process real-time video streams. Wankdhede et al. (2018) assessed the system’s sophistication through a series of shoplifting test cases. Their malicious attempts to steal items failed, proving that real-time video monitoring is an effective real-world tool. However, several critical changes are necessary before engineers can transpose the solution directly into a person’s home.

Action tracking systems require sophisticated machine learning algorithms that classify spatiotemporal movements. They typically follow a process that collects RGB+D (Color and Depth) camera frames and decodes them into 3-D space (Litomisky, 2012). Next, a series of filters must crop, resize, and align the principal subject. These steps are necessary because people can freely move around the room. Third, a feature extract process outputs matrices representing the body’s location, pose, and orientation. Finally, those tensors flow into Long-Term Short-Term (LTSM) algorithms that classify the movement into gestures.

Many general-purpose gesture detection libraries already exist for behaviors such as sporting events and other high-energy actions. However, daily indoor activity tends to be more subtle and nuanced (Das et al., 2019). This discrepancy limits reusability and creates the need for purpose-built training sets that sufficiently cover patient-specific actions. Furthermore, these custom datasets necessitate vast quantities of examples with accurate labels, or the machine learning model will produce unreliable results. As a general rule of thumb, each model parameter needs at least ten examples to avoid overfitting (Snee, 2015). Gesture models can quickly explode into hundreds or thousands of parameters which represents the 3-D space plus time. Mechanisms exist for accelerating the process of building custom datasets (e.g., transfer learning). However, this is still an open research topic.

## Integrating IoT Systems

After the system decodes the subject’s actions, it needs to act upon that information. While some responses are relatively trivial (e.g., dispatch an ambulance), other reactions must utilize Cyber-Physical Systems (CPS) to cross the digital boundary. Building these integrations is challenging as it draws upon knowledge spanning networking, sensors, embedded systems, and related concepts. Furthermore, CPS devices lack standardization which impacts component reusability.

Elloumi et al. (2020) propose a Smarthouse Operating System (SOS) that provides core services such as identity management, system state, and message routing. Their blueprint also outlines several automation application profiles (e.g., heat management and fire detection). These capabilities enable developers to focus on their integrations value differentiation versus writing tedious generic code. Finally, the authors demonstrate the effectiveness of this approach using a CNC (Computer Numerical Control) machine to print a replica house (2 cubic feet).

## Enhancing Security

It can be helpful to think of identity within IoT as a profile of historical choices, stated preferences, user roles, and known associations (Wachter, 2018). When the device understands the user’s profile, the experience can be customized and produce more accurate predictions. However, the payment for access to these inferences and decision processes comes from personal information, such as calendars, contacts, and routines (Mickens, 2018). This trade creates privacy concerns that can be subtle and can go unnoticed for some time.

The monitoring system will collect intimate knowledge of its assigned patients and even capture private communications. Therefore, mechanisms must exist for protecting this information as it flows between different compute domains (e.g., local versus cloud providers). Yang et al. (2021) propose a Secure MultiParty Computation (SMC) model that locally encrypts sensitive images (e.g., faces). Afterward, they use Cheon-Kim-Kim-Song Homomorphic Encryption (CKKS HE) to predict the encrypted payload remotely (e.g., Bob’s face). Ideally, the system minimizes the information that leaves the patient’s private network. However, when sensitive images must upload into the cloud, the system can leverage encryption strategies like CKKS HE.

## Healthcare and Cloud

Software that uses cloud resources gains agility, elasticity, instantaneous provisioning, and cost management constructs. However, some businesses are reluctant to trust these environments entirely due to security concerns (Ali et al., 2015). Alternatively, systems engineers can deploy hardware appliances that bring cloud aspects into the home or medical facility. For example, Abdulameer et al. (2020) propose implementing a small replica house similar to Elloumi. Their solution uses various wearable device sensors, Arduino micro-controllers, and one Raspberry-PI. In addition, users can check their vitals and other health KPIs (Key Performance Indicators) through a web portal. Therefore, provisioning small on-premise appliances makes sense and handles scenarios such as failures at the ISP (Internet Service Provider). In addition, other benefits come from standardizing the control-plane versus assuming the patient’s personal computer is compatible and Always-On Always Connected (AoAC).

Yi & Feng (2021) recently proposed a complete video-based injury rehabilitation solution that includes support for CPS and wearables. The authors leverage Carnegie Mellon University’s Open Pose library to map skeletal structures within images. Then, they publish this information and various sensor readings (e.g., smoke detectors) into a secure private cloud. Unlike Toyota (2019), the authors use Dynamic Time Warping (DTW) to compare and categorize patients’ movements. Researchers use DTW to normalize time series and avoid discrepancies from action speeds (e.g., raising one’s hand within two versus four seconds).

Chen, Saiki & Nakamura (2020) state that monitoring low-insensitive slow physical movements is challenging. These issues arise because training data is not broadly available due to researchers focusing on fast-paced sporting video by default. Their study uses PoseNet to track skeletal movements and predict activities. Additionally, they compare the resources necessary for Raspberry PI and a desktop computer to make those predictions. The researchers assess the model’s accuracy in terms of delta changes in the bounding box. This approach is unique and comes with several limitations. For instance, a person laying on the couch versus standing will have different dimensions. While this methodology is not directly usable, it provides another measurement strategy.

Chaing et al. (2011) propose a Uniform Markup Language (UML) model for collecting health care metadata from video sources. Their solution focuses on physiological information, such as the patient’s movements. The model also describes a storage structure for persisting the recordings. There are specific aspects from this study that are reusable. For instance, the authors propose service interfaces to several patient monitoring components (e.g., heart rate and oxygen levels). However, the video monitoring system is critically dependent on wearable technologies. This requirement makes the solution loosely relevant to the ESHOS project.

Nugroho, Harmanto & Al-Absi (2018) propose a deep learning model to assess a patients’ pain level. Their solution uses facial expressions from fourteen people that train both OpenFace and FaceNet topologies. The researchers claim that they can predict with 93% accuracy if a person is currently experiencing pain. This capability is helpful within home monitoring systems as a mechanism to assess medication levels. The ESHOS solution could introduce similar capabilities to improve patient care. For instance, the system could emit a pain frequency Key Performance Indicator (KPI) as part of the patient’s scorecard. The scorecard would then grant reassurance to family members that sufficient attention is available.