Chapter 3: Research Method

Nate Bachmeier

DIS-9902: Dissertation Proposal

May 28, 2022

Northcentral University

# Chapter 3: Research Method

*Write an introduction and chapter outcomes here.*

## Statement of the Problem

The problem to be addressed in this study is implementing a quality assurance process for an autonomous assistant to elderly and special needs care. Multiple industry-wide trends create the need for this technology. First, the number of practicing nurses has declined for several years (Kim & Kim, 2021). This labor shortage increases hiring and employee retention costs that the patients and welfare programs must cover. The funding gap is a global problem that does not impact all communities equally. For instance, in South Africa, rural special needs communities have 57% fewer nursing visits than their urban neighbors (Besada, 2020). Newly industrialized economies like Taiwan, South Korea, Thailand, and Malaysia are experiencing challenges maintaining their long-term care programs due to growing costs (Phua, 2021). Domestic programs like Veterans Health Administration (VHA) and Medicare are not immune to these economic limits (Lei et al., 2021). Businesses and governments must control these costs and replace human labor with less expensive automation processes.

Implementing and verifying those processes comes with a high barrier to entry, precisely due to personal privacy concerns, logistical complexity, ethical & cultural considerations, and procurement & configuration overhead. For example, a recent study shows that 95% of Pakistani versus 50% of New Zealand patients refuse to share a severe medical concern outside their primary care physician (Shirazi & Shekhani, 2021). Researchers create frameworks to mitigate these privacy concerns (e.g., redaction), though these procedures are challenging in practice (Blackhurn, 2021). Beyond human and process issues are technical complexities in configuring prototype autonomous assistants. It requires multiple domain specializations like computer networking, embedded technologies, AI/ML, and distributed computing (Tun, Madanian, & Mirza, 2021). Each cross-cutting concern adds complexity and reduces the probability that small teams can successfully provision their test environment. Furthermore, those difficulties limit other researchers from reproducing the results. These factors slow innovation and restrict the value researchers can contribute to the body of knowledge.

## Purpose of the Study

This constructive research design study aims to propose a research process that divorces privacy and safety concerns from investigating autonomous assistants in elderly and special needs care. It aims to deliver this capability by utilizing humanoid constructs within a realistic physics simulation process like PhysX or Gazebo (Bipin, 2018; Unreal, 2021). These engines support replaying specific MoCAP human behaviors under varying character properties such as weight, flexibility, and dexterity. Next, positioning virtual cameras, instruments, and devices within the virtual world enables researchers to collect their experimentation data. Lastly, the automation can modify the environment using programmable interfaces such as raising the alarm or applying other mitigations.

Hemodialysis (HD) patients have a high risk of falling and becoming injured (Shirai et al., 2021). This situation negatively impacts their quality of life by either remaining in bed or requiring more medical resources. The study explores this use case by virtualizing the HD patients and monitoring them with an AI/ML CV process to collect metadata and predict a fall in advance. Human trials prioritize safety, creating challenges to study metadata properties like floor slickness and character overexertion (Aihara et al., 2021). In contrast, humanoids are well-suited for these experiments. Furthermore, the lack of privacy concerns simplifies the video collection in bathrooms and showers.

Robot operating systems (ROS) and similar toolchains support generating dozens of floor plans and filling them with furniture (Bipin, 2018; A.W.S. RoboMaker, 2021). These services streamline experimentation, allowing the research to focus on the patient requirements versus simulation infrastructure. The study will use these capabilities to verify the AI/ML CV process across a reproducible gradient of character properties (e.g., weight from 80 to 500 lbs and age between 30 to 120 years).

## Research Methodology and Design

Design science is a research methodology that creates and uses purposeful artifacts to study a phenomenon (Hevner et al., 2004). 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 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.

### Study appropriateness

This methodology is appropriate *because…*

### Alternative methodologies

Quality research begins with a well-defined set of questions, such as ‘can an autonomous vehicle safely navigate city streets?’ Next, the researcher needs a plan to answer the question by collecting evidence and observations. Executing that plan requires a collection of quantitive and qualitative methods. Each of these methods is a tool with its inherent strengths and weaknesses (Jason & Glenwick, 2016). These attributes necessitate researchers to understand when a hammer is more appropriate than a screwdriver (see Table 1). Many people erroneously believe that quantitative methods are superior to qualitative alternatives (McCusker & Gunaydin, 2015; Creswell, 2014; Jason & Glenwick, 2016). This naïve perspective incorrectly assumes that a hammer is always the right tool. When researchers treat screws like nails, it results in erroneous publication claims.

Table 1: Research Approaches

|  |  |  |
| --- | --- | --- |
| Approach | Description | Example Use Case |
| Quantitative | Statistical modeling of a scenario | * Estimate the probability of an event * Stating a broad generalization * Cause and effect analysis |
| Qualitative | Non-numerical representation of a scenario | * Open-ended surveys * Exploration of needs * Investigating a local issue |
| Mixed-Method | Combination of both quantitative and qualitative | * Examining the breadth and depth of a topic * Examining a scientific idea and then mapping it to use cases |

Consider the difference when the vehicle study’s objective is (a) to identify safety requirements versus (b) modeling the limitations of the braking system. Under (a), qualitative methods best support the open exploratory nature of the problem. With (b), the answer needs a quantitative method that describes the relationship of multiple variables, such as the car’s speed and the number of objects on the road. However, a more comprehensive study could answer both (a) and (b) by uncovering the importance of braking enhancements and then describing the limitations in greater detail.

*Now let’s take that analogy back to the original problem statement. There should also be a page(?) that describes how each of these studies could look before saying that these are ultimately different problems than the goal here… and that goal is to study neural networks in a quasi-realistic use-case.*

## Population and Sample

For an experiment to be successful, it needs to have sufficient *power* to measure the *effect* in question. Several knobs feed into the power of an experiment, such as relaxing the confidence interval, using parametric statistics, converting to a one-tail model, increasing the samples, or adjusting the sensitivity (Donovan, 2016). Choosing which value to tweak and optimize is scenario-specific and can be somewhat of an art form.

### Determining Power

*Insert data about choosing the power level here*. Given the relatively small sample count, adjusting the confidence intervals to meet acceptable power requirements might be necessary. Another option might be to reduce the number of racial categories, from nationalities to three groups. These data tweaks might detect high-level trends that future research could tease further.

### Determining Effect

Effect size measures the strength of a phenomenon (Donovan, 2016). While calculating the difference between the two distributions is relatively straightforward, it can be difficult to predict ahead of time. This bittersweet relationship introduces challenges when determining the appropriate sample size. One potential solution is to use an iterative sequential sampling policy instead of a fixed size upfront (García-Pérez, 2012). In this situation, *insert data about what this means for this situation* and variable selection. While this small group would have a reasonably low confidence interval, it could qualitatively hint at the overall sample size needing to be minor, medium, or large. There are potential risks that the random-initial sample produces an invalid seed in the study.

### Potential Sample Sizes

Despite the effect size being unknown potential, it is possible to determine the range of sample sizes for the experiment (see Table 2). G\*Power version 3.1.9.7 projects that t-tests of the “difference between two independent means (two groups)” for a one-tail model will need somewhere to *insert actual values*. Since *insert available values* examples, there should be sufficient coverage assuming the specific measurements are kept simple.

Table 2: Sample Sizes

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Power | Effect Size | Confidence – 50% | Confidence – 80% | Confidence – 95% | Confidence – 99% |
| 70% Adequate | 0.20 – Small | 28 | 188 | 472 | 816 |
| 0.50 – Medium | 6 | 32 | 78 | 134 |
| 0.80 – Large | 4 | 14 | 32 | 54 |
| 95% Excellent | 0.20 – Small | 272 | 620 | 1084 | 1580 |
| 0.50 – Medium | 44 | 100 | 176 | 256 |
| 0.80 – Large | 18 | 40 | 70 | 102 |

### Acquiring the required sample

This study aims to demonstrate a research methodology for using humanoids in simulation processes to assess machine learning models. It presents an example of employing computer vision (CV) to detect falling patients. The research project will generate different humanoid configurations and have them perform MoCAP sequences. For instance, one experiment would provide a thirty-year-old actor that’s one hundred pounds (forty-five kilograms). Another one could have a sixty-year-old actor that weighs three hundred pounds. The simulation software will use these variables to influence movement speed and flexibility.

Using this approach is appropriate for the dissertation proposal methodology and design. It has several core strengths, such as avoiding a cumbersome human recruiting process and concerns that the selection procedure is unfair. This method examines the generalization and usefulness of the research technique. Furthermore, the experiments automated nature makes reproducing the results straightforward and economical. This design choice means that future researchers have sufficient information to replicate the study.

## Instrumentation

In a physics simulation process, humanoid actors perform behaviors in a highly controlled environment. This feature allows the study always to know the current world state and quickly assess any CV model prediction’s accuracy.

*There should be another 2-paragraphs with examples of this idea here.*

*Checklist:*

*Describe the instruments (e.g., tests, questionnaires, observation protocols) that will be (proposal) or were (manuscript) used, including information on their origin and evidence of their reliability and validity. OR as applicable, describe the materials to be used (e.g., lesson plans for interventions, webinars, or archived data, etc.).*

*Describe in detail any field testing or pilot testing of instruments to include their results and any subsequent modifications.*

*If instruments or materials are used that were developed by another researcher, include evidence in the appendix that permission was granted to use the instrument(s) and/or material(s) and refer to that fact and the appendix in this section.*

## Study Procedures

The research project aims to build a CV model that can accurately predict human activity recognition (HAR). Model training will initialize a random experiment configuration and perform an appropriate MoCAP sequence. During the performance, a virtual camera will collect changes in joint positionings. This delta stream will serve as input feature parameters to the classification process (e.g., sitting versus falling).

### Building the Model

A distributed training service can horizontally scale and assess these different humanoid permutations in isolation. Amazon SageMaker offers these capabilities through its “bring your own container” design. Researchers essentially bundle custom automation and open-source tooling into a virtualized process. SageMaker uses public cloud resources like compute and storage to execute the experiment hundreds or thousands of times. It also integrates into TensorFlow 2 for collecting accuracy and performance metrics. These features reduce the complexity of building boilerplate instruments for many standard requirements.

Future researchers can replicate this experiment by deploying the same container images into their Amazon SageMaker and TensorFlow 2 environments. The humanoid automation will be versioned using GitHub. GitHub simplifies sharing open-source code and identifying specific point-in-time versions (called a commit SHA). Since those researchers can synchronize the repository to a particular commit and rerun the automation using industry-standard tooling, they have sufficient capabilities to reproduce the experiment.

### Implementing the Simulation Process

*Include 2-3 paragraphs that detail how the simulation will take place. Like, is it a Unity process with OpenGym? What open-source software will go into the container? How will that map into handling the research questions? This information is critical and doesn’t currently exist. Those gaps are likely another section in the Literature Review to define them.*

### Using the Model

*Data Analysis > Model Inference Analysis mentions a small cohort (co-workers) that could interact with the model endpoint. What exactly does that look like, and how would it work?*

## Data Analysis

There are two phases to implementing an AI/ML process: training the model and operationalizing the capability. The analysis must confirm that these phases meet acceptable quality standards. Additionally, it must succinctly address the research questions from chapter 1.

### Model Training Analysis

Tensor Flow 2 generates statistical information regarding the model training performance. These Key Performance Indicators (KPIs) describe the gradient convergance, model accuracy, and various troubleshooting metrics. An analysis must confirm that the training configuration occurs efficiently. Suppose the performance is substandard. In that case, this research plans to investigate the defects and reconfigure the training service (e.g., Amazon SageMaker). It is beyond this study’s scope to create “a perfect model” and only seeks to demonstrate the research technique’s viability. However, this study will validate that training is reliable and reproducible across positive and negative test cases.

### Model Inference Analysis

An analysis of the model inference must confirm that it is usable. This phase requires provisioning a model endpoint and posting experimental data. A simple approach could be using RGB+D cameras to record a small human group repeating the humanoid behaviors. There are several core advantages to this solution. First, it demonstrates bringing the simulation process into the real world. Next, these volunteers are readily available through work and social gatherings. It is beyond this study’s scope to “perfectly predict” every behavior. Instead, the goal is to collect and evaluate operationalizing the research technique.

### Addressing the research questions

*Repeat the research questions. Describe in 1-2 paragraphs what data is necessary and how it will be analyzed. Talk about statistical functions, their appropriateness, and their alignment with the hypothesis.*

Checklist:

~~Describe the strategies that will be (proposal) or were (manuscript) used to code and/or analyze the data, and any software that will be (proposal) or was (manuscript) used.~~

Ensure the data that will be (proposal) or were (manuscript) analyzed can be used to answer the research questions and/or test the hypotheses with the ultimate goal of addressing the identified problem.

Use proper terminology in association with each design/analysis (e.g., independent variable and dependent variable for an experimental design, predictor and criterion variables for regression).

For quantitative studies, describe the analysis that will be (proposal) or was (manuscript) used to test each hypothesis. Provide evidence the statistical test chosen is appropriate to test the hypotheses and the data meet the assumptions of the statistical tests.

For qualitative studies, describe how the data will be (proposal) or were (manuscript) processed and analyzed, including any triangulation efforts. Explain the role of the researcher.

## Assumptions

Research projects must be cognizant of the internal and external factors influencing their research. Making an assumptions inventory is essential to quality research because it communicates the implicit drivers in the design.

### Quality Tooling Exists

This study makes several assumptions about the current industry state. It assumes that mainstream solutions like Amazon SageMaker, Robotic Operating System, Docker, OpenAI’s Gym, and Unity’s PhysX deliver the capabilities necessary to build the core artifacts. This situation would allow the experiment to focus on the AI/ML components, not rewriting boilerplate infrastructure. Suppose the toolsets haven’t matured to an acceptable level. In that case, the study will simplify the training subsystem.

### Adequate Funding Exists

The current plan also assumes access to a highly discounted rate for cloud computing resources. Amazon Web Services (AWS) has several programs for aiding researchers, like AWS Cloud Credit for Research and AWS Educate. Presently this study has funding through one or more of these programs and can pursue the entire project’s scope. Suppose that Amazon discontinued funding. In that case, the study would reduce the scale and focus on fewer test cases and humanoid configuration combinations.

### CV Models Can Predict HAR

This dissertation aims to demonstrate a research technique using computer vision to predict human activity recognition. Several researchers are documenting their successful experiments within the field. However, this is a state-of-the-art topic, and the underlying example might not work. The study aims to communicate the open problems and potential next steps in this case. While this study makes every effort to mitigate critical blocks efficiently, it is beyond the scope of the core research.

### Simulation Processes and MoCAP are Compatible

There is an assumption that open-source motion capture (MoCAP) files are compatible with industry-standard physics simulation processes. The test cases aim to use virtual cameras to capture this information in 3-D open worlds. Suppose it is not possible to reuse that footage. In that case, the study can flatten the MoCAP to 2-D and present the findings. This approach is not as impressive but would complete the dissertation requirements.

## Limitations

Limitations are internal and external factors that *implicitly* restrict the study from exploring all aspects of the problem.

### Range of Motion

This study aims to build a HAR classification model that supports a predefined set of activities. These limitations exist due to challenges in finding sufficient example data. In this case, expanding the sample to contain open-source repositories will become necessary. These repositories could include YouTube, among other sites.

## Delimitations

Deliminiations are internal and external factors that *explicitly* restrict the study from exploring all aspects of the problem.

### Humanoid Constraints

Humanoid actors initialize with a configuration that controls their mechanical movement. There are virtually unlimited permutations for these characters and their weight, height, dexterity, and flexibility, among other properties. The distributed training process must set value bounds to learn the problem space efficiently. For instance, there’s only one person over a 635KG weight (Guinness World Records, 2022). Therefore, it does not make sense for test cases to exceed this extreme limit. Similar practical constraints also exist for other properties. It is beyond the scope and budget of this study to examine outliers.

## Ethical Assurances

Northcentral University’s Institutional Review Board (IRB) must issue a statement covering ethical concerns, privacy violations, or undue harm risks.

### Human Subject Concerns

This study uses humanoids in a physics simulation process as a research technique that mitigates ethical concerns and personal privacy risks. Since a humanoid is a virtual construct, it intentionally and explicitly divorces any moral hazards. Furthermore, the simulation has no right or assumption to privacy, as it does not exist in the real world. To verify the model training, a small cohort of volunteers will re-enact safe behaviors that do not risk personal privacy or harm. For instance, the falling behavior can be onto a padded surface.

### Secure Data Storage

*Include a paragraph that all data will be encrypted at rest and in transit. It will use AAA and all that good stuff. Maybe there’s something in CyberSec course to snag into this section?*

### Researcher Concerns

*Include a paragraph that talks about past experiences that might taint the experiment or skew its correctness.*

Checklist:

~~Confirm in a statement the study will (proposal) or did (manuscript) receive approval from Northcentral University’s Institutional Review Board (IRB) before data collection~~.

~~If the risk to participants is greater than minimal, discuss the relevant ethical issues and how they will be (proposal) or were (manuscript) addressed.~~

~~Describe how confidentiality or anonymity will be (proposal) or was (manuscript) achieved.~~

~~Identify how the data will be (proposal) or were (manuscript) securely stored in accordance with IRB requirements.~~

~~Describe the role of the researcher in the study. Discuss relevant issues, including biases and personal and professional experiences with the topic, problem, or context. Present the strategies that will be (proposal) or were (manuscript) used to prevent these biases and experiences from influencing the analysis or findings.~~

In the dissertation manuscript only, include the IRB approval letter in an appendix.

## Summary

Begin writing here…

Checklist:

Summarize the key points presented in the chapter.

Logically lead the reader to the next chapter on the study's findings.

# References

Adashi, E., Walters, L., & Menikoff, J. (2018). The Belmont Report at 40: Reckoning With Time. *American Journal of Public Health, 108*(10), 1345-1348. Retrieved from https://search-ebscohost-com.proxy1.ncu.edu/login.aspx?direct=true&db=s3h&AN=131743373&site=eds-live

Aihara, S., Kitamura, S., Dogan, M., Sakata, S., Kondo, K., & Otaka, Y. (2021). Patients' thoughts on their falls in a rehabilitation hospital: a qualitative study of patients with stroke. *BMC Geriatrics, 21*(1), 1-12. Retrieved from https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,sso&db=edb&AN=154196006&site=eds-live&scope=site

Besada, D. E. (2020). Resource requirements for community-based care in rural, deep-rural and peri-urban communities in South Africa. *PLoS ONE, 15*(1), 1-19. doi:https://doi.org/10.1371/journal.pone.0218682

Bipin, K. (2018). *Robot Operating System Cookbook.* Packet Publishing.

BJS. (2015, October). *Contacts between police and the public*. Retrieved from Bureau of Justice Statistics: https://www.bjs.gov/content/pub/pdf/cpp15\_sum.pdf

Blackhurn, B. (2021). Sensitive situations in a nurse residency program: balancing confidentiality with meaningful solutions. *Journal for Nurses in Professional Development, 37*(3), 185-187. doi:10.1097/NND.0000000000000694

Bryar, C., & Carr, B. (2021). *Working Backwards: Insights, Stories, and Secrets from Inside Amazon.*

Burr, V. (2015). *Social Constructionism.* Routledge.

Creswell, J. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches.* Thousand Oaks, CA: Sage Publishing, Inc.

Donovan, C. (2016, August 30). *Power and effect size*. Retrieved from YouTube: https://www.youtube.com/watch?v=9LVD9oLg1A0

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

Gergen, K. (2010). *Social Constructionist Ideas, Theory and Practice*. (The Taos Institute) Retrieved from Vimeo: https://vimeo.com/15676699

Hevner, A., March, S., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS. Quarterly, 28*(1), 75-105. doi:10.2307/25148625

Jason, L., & Glenwick, D. (2016). *Handbook of methodological approaches to community-based research : qualitative, quantitative, and mixed methods .* Oxford University Press.

KBP. (2020, August 11). *Police shootings database*. Retrieved from Killed By Police: https://killedbypolice.net/

Kim, J., & Kim, S. (2021). The determinants of caregiver use and its costs for elderly inpatients in Korea. *BMC. Health Services Research, 21*(631), 1-10. doi:https://doi.org/10.1186/s12913-021-06677-w

Lei, L., Intrator, O., Conwell, Y., Fortinsky, R., & Cai, S. (2021). Continuity of care and health care cost among community-dwelling older adult veterans living with dementia. *Health Services Research, 56*(3), 378. Retrieved from https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,sso&db=edsgao&AN=edsgcl.667587895&site=eds-live&scope=site

McCusker, K., & Gunaydin, S. (2015, October). Research using qualitative, quantitative or mixed methods and choice based on the research. *Perfusion, 30*(7), 537-542. doi:10.1177/0267659114559116

Office of the Secretary. (1979). *The Belmont Report*. Retrieved from US Department of Health and Human Services: https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c\_FINAL.pdf

Owen, C. (2017, November 8). *A Theorical Hands-on introduction to Fouculdian Discourse Analysis*. Retrieved from YouTube: https://www.youtube.com/watch?v=6I6b3ePAZ5M

Peffers, K., Tuunanen, T., Rothenberger, M., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45-77. doi:10.2753/MIS0742-1222240302

Phua, K. H. (2021). Ageing in Asia: beyond the Astana declaration towards financing long-term care for all. *International Journal of Health Policy and Management, 10*(1), 32-36. doi:https://doi.org/10.34172/ijhpm.2020.15

Roberts, C. (2021). The Ethics of Research That May Disadvantage Others. *Ethics & Human Research, 43*(1), 2-16. doi:10.1002/eahr.500074

Shirai, N., Yamamoto, S., Osawa, Y., Tsubaki, A., Morishita, S., Nitami, S., & Narita, I. (2021). Fear of falling and physical activity in hemodialysis patients. *Renal Replacement Therapy, 7*(1). doi:10.1186/s41100-021-00383-3

Shirazi, B., & Shekhani, S. (2021). Patient’s expectations of privacy and confidentiality in Pakistan. *The Journal of the Pakistan Medical Association, 71*(2A), 537-539. doi:https://doi.org/10.47391/JPMA.888

Tan, Z. (2021). Ethics Events and Conditions of Possibility. *Business Ethics Quarterly, 31*(1), 106-137. Retrieved from https://search-ebscohost-com.proxy1.ncu.edu/login.aspx?direct=true&db=edb&AN=147839336&site=eds-live

Tun, S., Madanian, S., & Mirza, F. (2021). Internet of things (IoT) applications for elderly care: a reflective review. *Aging Clinical & Experimental Research, 33*(4), 855-867. doi:10.1007/s40520-020-01545-9

Unreal Engine. (2022). *Unreal Engine Product*. Retrieved from Unreal Engine: https://www.unrealengine.com/en-US/unreal-engine-5