Applied Human Factors and Ergonomics (AHFE2025), Vol. 199, 2025, 1057–1066
https://doi.org/10.54941/ahfe1006918

Leveraging AI and Multivariate Analysis to Convert Product Requirements Into Product Specifications

Zoe Marazita[1], Matt Parkinson[1,2], and Jessica Menold[1,3]

[1]Department of Mechanical Engineering, The Pennsylvania State University, State College, PA 16802, USA
[2]School of Engineering Design and Innovation, The Pennsylvania State University, State College, PA 16802, USA
[3]Harold and Inge Marcus Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, State College, PA 16802, USA

ABSTRACT

The objective of this work is to enable designers to easily determine the optimal product dimensions to accommodate diverse user populations. Artificial intelligence (AI) is integrated with data on a population's body size and shape (anthropometry) and a custom analysis function to convert natural-language design requirements to technical design specifications. This leverages a strength of AI to help designers overcome the challenges of unfamiliarity with anthropometric terminology. Simultaneously, it mitigates some limitations of AI by performing the analysis in an environment specifically designed for this task. Ultimately, it allows human factors engineers and ergonomists to easily explore design trade-offs in a multivariate design space.

Keywords: Human-centred design, Ergonomics, Multivariate design, Artificial intelligence, Anthropometry

INTRODUCTION

In human factors and ergonomics, considering a wide range of anthropometry is essential for ensuring that products are usable and safe (Pheasant and Haslegrave, 2006; Human Factors and Ergonomics Society, 2004). However, the process of determining the dimensions of a product to fit a target percentage of the population is often challenging for designers, particularly in the context of *multivariate design*, where multiple dimensions must be addressed simultaneously (Hudson et al., 1998; Porter et al., 2004). Unlike univariate design, which focuses on optimizing a single dimension at a time, multivariate design involves balancing multiple variables, each affecting the accommodation in different ways. While some tools for assisting in multivariate design exist, such as the Multivariate Anthropometry Testing Tool from the Human Factors and Ergonomics Society (HFES, 2020), they necessitate an understanding of anthropometry (body dimensions). Designers must identify the relevant anthropometric measures for each product variable and understand each measure's relation to the product.

Artificial intelligence (AI) has great potential in making human-centred design more accessible for designers (de Winter et al., 2025; Petrat, 2021). Current Generative Pre-trained Transformer (GPT) models can identify public datasets, such as the ANSUR II data, which contain detailed anthropometry from military personnel (Gordon et al., 2014). The models are also able to match product dimensions to relevant anthropometric measures, inferring the relationships between a design variable and the associated anthropometry. However, the existing models struggle in two critical ways. First, they fail to reliably extract information from online datasets and often report incorrect data for design recommendations. Second, they are unable to conduct multivariate analyses. When asked to size a product in more than one dimension, the current models will report a series of independent univariate solutions, which is known to overestimate performance (da Silva et al., 2020; Vega et al., 2021; Parkinson and Reed, 2010).

To address these challenges, this work introduces the incorporation of function calling in GPT to improve multivariate accommodation analysis. *Function calling* is a capability of GPT models that enables them to conduct custom analysis beyond the model's default reasoning (Wang et al., 2025). In this case, it allows the GPT to trigger a custom backend process that directly accesses ANSUR II anthropometric data and accurately computes accommodation based on multiple dimensions. Because the function correctly performs multivariate analyses, the GPT can provide several design recommendations that meet the overall target, allowing the designer to select the most appropriate one. By allowing the GPT to infer user intent and adjust function arguments, this approach overcomes key limitations of existing tools and GPT models, enabling more efficient and accurate design solutions. For the present work, variables related to seat design (seat pan width, depth, and height) are used as a specialized case study for prompting GPT models with multivariate design problems. These design variables are labelled in Figure 1. An individual is considered "accommodated" when their seated hip breadth is smaller than seat width, their buttock-popliteal length is larger than seat depth, and their popliteal height is larger than seat height (Sydor and Hitka, 2023). This is a simplification of the requirements outlined in HFES 100–2007 (HFES, 2007).

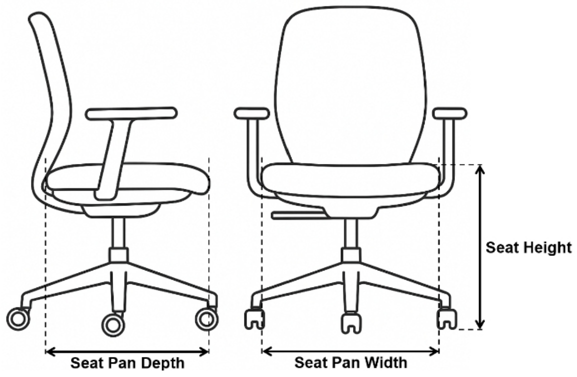

Figure 1: Diagram of chair parameters used as a case study for AI-driven multivariate design recommendations.

Univariate Design

Univariate design in ergonomics refers to the process of analysing a single design variable or anthropometric measure at a time (Dianat et al., 2018). This process often involves selecting design parameters based on the percentiles of a specific anthropometric measure, such as the 5^{th} or 95^{th} percentile of stature (Robinette and McConville, 1981). While this approach can be sufficient for designing products with only one variable, most ergonomic problems have multiple variables, each related to one or more anthropometric measures. Use of univariate analysis in these situations fails to accurately represent the needs of a diverse user population (Dianat et al., 2018).

Multivariate Design

Multivariate design refers to the sizing of multiple product dimensions and considering their interactions with corresponding anthropometric variables simultaneously (Hudson et al., 1998). This approach is essential for designing inclusive products that will accommodate a wide range of body shapes and sizes (Porter et al., 2004; Hsiao, 2013; Guan et al., 2012). For example, in the design of task seating, variables such as seat pan width, depth, and height must be analysed simultaneously to estimate resulting accommodation. If each dimension is sized independently to fit 95% of the population, the resulting population-level accommodation will be substantially less than the target.

Despite the importance of using multivariate analysis to meet accommodation targets, multivariate design methods are often misunderstood due to their complexity and the lack of accessible tools (Porter et al., 2004). Designers must be able to identify relevant anthropometric variables, understand their relationships to design variables, and have access to the detailed data (not just summary statistics) to perform the necessary calculations. Artificial intelligence (AI), specifically natural language models such as GPT, offers a promising resource for this design process (Petrat, 2021). However, current models are limited in this domain. They frequently misinterpret anthropometric datasets (e.g., ANSUR II military data, and NHANES health data for the U.S. population) and cannot accurately perform multivariate analysis and instead default to univariate approximations.

METHODS

To evaluate the performance of current GPT models in solving multivariate design problems, a baseline assessment was conducted using ChatGPT 4 (OpenAI, 2023). This baseline highlights critical limitations in the model's multivariate analysis and data extraction capabilities. To address these shortcomings, a custom multivariate analysis function was implemented using GPT function-calling. This allows natural language prompts to be translated into accurate, data-driven accommodation estimates.

Evaluating GPT's Multivariate Design Performance

We began by prompting ChatGPT with a multivariate design problem related to seat sizing:

"I am designing a chair. How wide and deep should the seat pan be so that 90% of the US population fits?"

The GPT responded by calculating accommodation based on percentiles of seated hip breadth and buttock-popliteal length from the ANSUR II dataset using the following logic:

1. Seat width should be greater than or equal to the 90^{th} percentile of U.S. seated hip breadth.
2. Seat depth should be less than or equal to the 10^{th} percentile of U.S. buttock-popliteal length.

This demonstrates the model's ability to relate design variables to relevant anthropometric measures. However, several critical limitations were observed. First, the GPT recommended a seat width of 430 mm and seat depth of 430 mm, claiming that they correspond to 90^{th} percentile seated hip breadth and 10^{th} percentile buttock-popliteal length from ANSUR II data. This solution not only uses military data to generalize recommendations for the US population, but also does not clarify whether male and/or female ANSUR II data were used. In reality, the 50:50 male:female ANSUR II data has a 90^{th} percentile seated hip breadth of approximately 428 mm and a 10^{th} percentile buttock-popliteal length of approximately 461 mm. This error suggests that the model may have fabricated values that are not present in the dataset. Additionally, the model evaluated each design variable independently, leading to an overestimation of accommodation. When evaluated with the ANSUR II dataset using multivariate analysis, the GPT's recommended dimensions accommodate only 88.3% of the population, falling short of the 90% target.

Developing Improved Multivariate Design Capabilities for GPT

To improve the performance of current GPT multivariate analysis, we developed a tool for natural language-driven design that leverages OpenAI's GPT-4 model (OpenAI, 2023). This tool integrates GPT-4 with a custom backend function in Python (Van and Drake, 1995) that correctly recommends product sizes using the ANSUR II dataset for design problems with one, two, or three variables.

Function-Calling Implementation

When a user inputs a natural language prompt (e.g. "What should a chair's seat width and height be to fit 90% of adults?"), the tool sends the prompt to GPT-4 along with a custom system prompt that instructs the model to call a predefined accommodation analysis function. The function requires the following structured argument inputs which are inferred by GPT-4 from the user's prompt:

1. product: (e.g. "chair seat"'),
2. target accommodation level: a value between 0 and 1 (e.g. 0.90),
3. product variables: a list of one to three design variables (e.g. "seat height"), each specifying relevant anthropometric measure (mapped to

ANSUR II variables), and the relation between the variable and measure (either “>” or “<”),

4. range of values ([min, max, step]) for evaluation.

The system prompt helps to ensure the GPT selects appropriate anthropometric measures from ANSUR II data, applies the correct logical operators, and recommends realistic ranges covering approximately the 1st to 99th percentile of the ANSUR II population to ensure thorough evaluation.

Parsing and Range Selection

After the GPT calls the backend function, the arguments are parsed and validated so that each anthropometric measure is matched to a column in the ANSUR II dataset using fuzzy string matching. If no match is found, the process stops and returns an error. If a range is not sent to the multivariate analysis function, the tool calculates the range based on empirical percentiles from the ANSUR II column and selects a consistent step to evaluate a range of values (always between 1–5 mm).

Anthropometric Accommodation Analysis

The backend function uses the ANSUR II male and female datasets with assigned statistical weights for male and female samples to represent a 50% male, 50% female population. The accommodation logic for a given combination of product dimensions is calculated using the following equation:

$$\text{Accommodation} = \frac{\sum \text{indivivuals accommodated in all dimensions (weighted)}}{\text{Total weighted sample size}}$$

To determine the percent accommodated in each size combination, every individual is evaluated against the specific design criteria (e.g. seat width > seated hip breadth). The combined weighted sum of individuals who fit in the combination is divided by the total weight. For 1- and 2-variable problems, the design space is evaluated over a full grid of sizes and combinations. For 3-variable problems, random sampling is used to reduce computational complexity.

Data Visualization and User Interface

The entire tool is deployed as an interactive web application using Streamlit (Streamlit, 2025). This allows users to enter their natural language design prompts and an OpenAI API key. They are also able to view a table of the five sizing combinations that are closest to their desired accommodation level and explore visual design tools displaying a wide range of accommodation results.

The generated plot adapts based on the number of design variables. A single-variable prompt results in a cumulative distribution function (CDF), showing the proportion of the population accommodated across the range of sizes. For two-variable prompts, a heat map is generated with each grid cell reflecting the percentage of the population accommodated from the combination. Finally, for three variable design prompts, the variable with greatest variance is selected to be frozen, and sliced into nine evenly spaced levels. Then, heat maps are generated for each of the nine levels, showing accommodation across combinations of all three variables in "small multiples" plots.

RESULTS

By integrating function-calling with OpenAI's GPT, the tool is able to accurately extract ANSUR II anthropometric data, recommend design solutions to 1, 2, and 3 variable problems, and generate visual design tools for analysing the trade-offs of different sizing configurations.

Multivariate Analysis Performance

When prompted with a design problem, the tool iterates through a range of product size combinations and reports five size configurations that achieve an accommodation level closest to the user's target. To compare the design tool's performance with the baseline from ChatGPT, the tool was given the same original prompt:

"I am designing a chair. How wide and deep should the seat pan be so that 90% of the US population fits."

The GPT tool responded with five sizing recommendations as well as their resulting level of accommodation within the ANSUR II database (Table 1). One limitation of this current implementation is that it uses military data to generalize for the U.S. population.

Table 1: The top five seat pan depths and widths recommended by the design tool to reach a target of 90% accommodation. The corresponding percent accommodated in the ANSUR II data are also listed.

Seat Depth (mm)	Seat Width (mm)	Percent Accommodated
410	432	90.00%
411	432	90.00%
432	434	90.00%
457	479	90.00%
455	464	90.01%

Visual Design Tools

The multivariate analysis tool was also able to generate accurate accommodation plots that adapt based on how many design variables are entered by the user. When the user enters a design problem with one variable, a cumulative density function (CDF) is generated with accommodation as a function of the design variable values. For example, Figure 2 displays a CDF for varying seat widths. As expected, when seat width increases, the percent of accommodated individuals increases. The target percent accommodation from the user's prompt is also displayed on this plot.

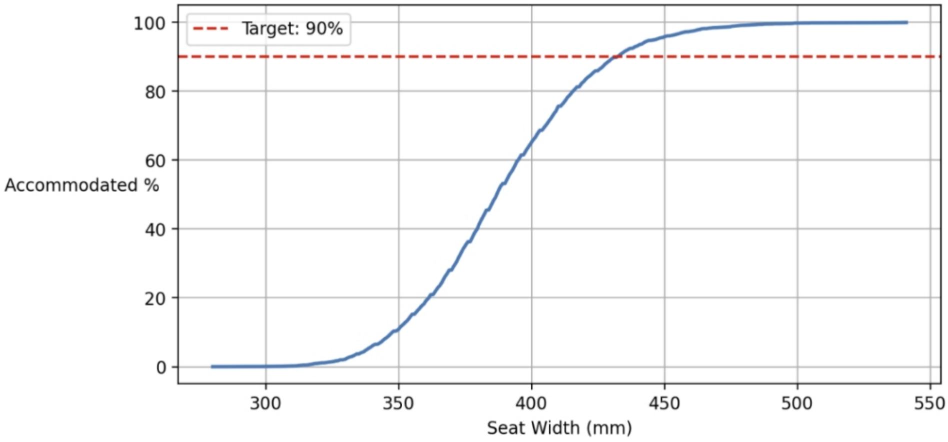

Figure 2: Cumulative density function of accommodation as a function of seat pan width (mm), generated using the updated GPT design tool.

When the tool is prompted with two design variables, a heat map is generated with the variables on the x and y-axes and a colour scale corresponding to the percent of individuals accommodated. In Figure 3, a heat map was generated from a design prompt for seat width and depth sizes. In this plot, higher levels of accommodation correspond with smaller seat depth and larger seat width, which is the expected result. The plots generated by the multivariate design tool use a distinguishable colour scheme for visualizing changes in accommodation.

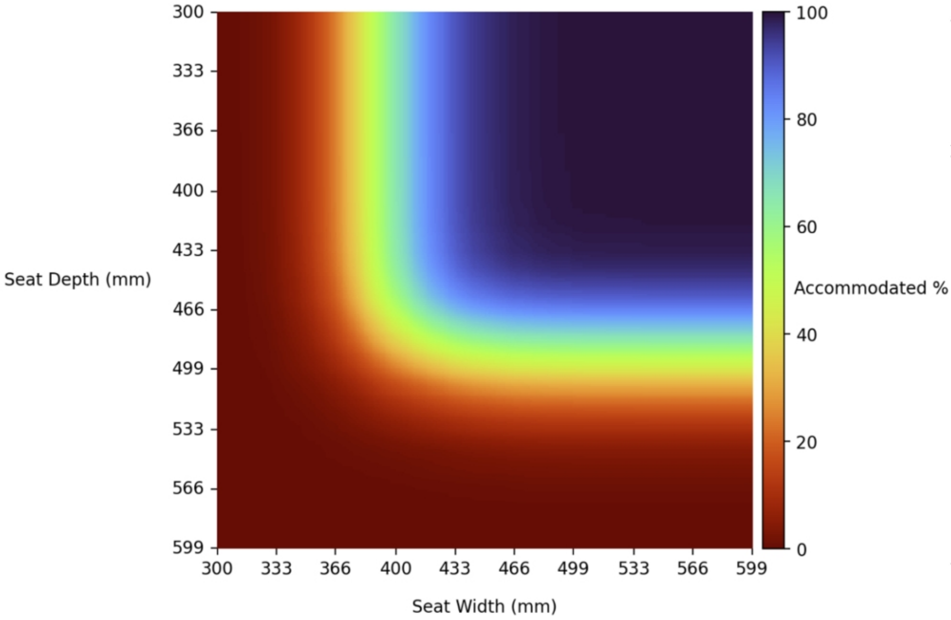

Figure 3: Heat map of seat pan width and depth (mm) accommodation and the resulting level of accommodation generated by the updated GPT design tool.

When the design prompt contains three design variables, a small multiples plot is generated containing nine heat maps. For each heat map, one variable is frozen across nine levels and the x and y-axes correspond to the other two design variables. Accommodation is represented by a colour scale and calculated from the combination of all three variables. In Figure 4, the nine heat maps represent varying seat widths while seat height and depth are varied within each plot. As expected, accommodation increases as seat width and height increase, and seat depth decreases.

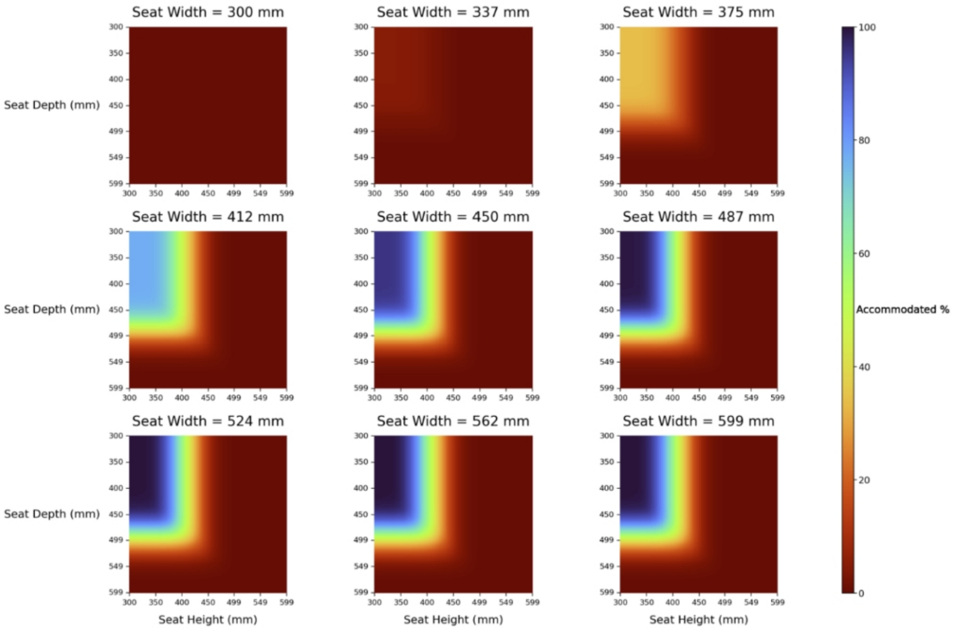

Figure 4: Heat maps of the resulting accommodation across different levels of seat width in terms of seat pan depth and height (mm), generated by the updated GPT design tool.

CONCLUSION

This work presents an approach to improving accessibility of multivariate anthropometric design by integrating GPT with function-calling capabilities. Sizing a product across multiple dimensions is often a complex task, requiring knowledge of anthropometric terminology and multivariate analysis methods (Hsiao, 2013). By allowing users to describe their design problems using natural language prompts, this tool helps to provide data-driven solutions to human-centred design problems.

Function calling enables GPT to overcome key limitations by offloading technical calculations to a backend process specifically designed for anthropometric analysis. This integration leverages the GPT model's strength in understanding natural language while mitigating its weaknesses in accessing data and multivariate logic.

The case study in this work focuses on seat design (seat pan width, depth, and height) and demonstrates the tool's ability to interpret user intent, identify relevant anthropometric relationships, and produce multiple design recommendations to meet a population-level accommodation target. The current ChatGPT responses to multivariate design prompts recommend one size configuration that overestimate the actual percentage of people accommodated. By providing multiple solutions that achieve the target accommodation level, the multivariate tool allows designers to explore trade-offs and select a sizing configuration that best suits their application. In addition to improved consistency with the calculation methods and data extraction, the tool is also able to generate useful design tools for visualizing accommodation across several combinations of each design variable.

Limitations and Future Work

While this work demonstrates a significant advancement for AI-driven ergonomic design, several improvements could be made. First, generalizing this tool for a wider range of products will require enhanced prompt interpretation and an expanded mapping of design variables to anthropometric measures. The use of retrieval-augmented generation (RAG) is a promising method for ensuring that GPT responses are grounded in reliable ergonomic references for improved consistency across design applications (Lewis et al., 2021). This also uses ANSUR II military data to make design recommendations for the U.S. population. This could be replaced by more representative anthropometric data for the U.S. to improve design recommendations.

Another challenge is model accessibility. This preliminary tool uses OpenAI's proprietary GPT models, which require paid API access. Open-source GPT models currently demonstrate reduced consistency in inferring anthropometric relationships, which would reduce the tool's reliability. Improving the performance of open-source alternatives could help democratize this tool for public use. Overall, this work demonstrates that combining AI and ergonomics principles through function-calling can create an intuitive, scalable tool for conducting multivariate design, allowing more designers to create inclusive, data-driven products and environments.

REFERENCES

da Silva, G. V., Zehner, G. F. & Hudson, J. A. (2020). "Comparison of univariate and multivariate anthropometric design requirements methods for flight deck design application." *Ergonomics*, 63, pp. 1133–1149. doi:10.1080/00140139.2020.1765029.

de Winter, J. C. F. & Eisma, Y. B. (2025). "Ergonomics and human factors: still fading—and why we need to embrace the AI revolution." *Ergonomics*. Advance online publication. doi:10.1080/00140139.2025.2509352.

Dianat, I., Molenbroek, J. & Castellucci, H. I. (2018). "A review of the methodology and applications of anthropometry in ergonomics and product design." *Ergonomics, 61(12), pp. 1696–1720. doi:10.1080/00140139.2018.1502817..*

Gordon, C. C., Blackwell, C. L., Bradtmiller, B., Parham, J. L., Barrientos, P., Paquette, S. P., Corner, B. D., Carson, J. M., Venezia, J. C., Rockwell, B. M., Mucher, M. & Kristensen, S. (2014). *2012 Anthropometric Survey of U. S. Army Personnel*. NATICK/TR-15/007. U. S. Army Natick RD&E Center.

Guan, J., Hsiao, H., Bradtmiller, B., Kau, T.-Y., Reed, M. R., Jahns, S. K., Loczi, J., Hardee, H. L. & Piamonte, D. P. T. (2012). "U. S. Truck Driver Anthropometric Study and Multivariate Anthropometric Models for Cab Designs." *Human Factors*, 54(5), pp. 849–871. doi:10.1177/0018720812442685.

Hsiao, H. (2013). "Anthropometric procedures for protective equipment sizing and design." *Human Factors*, 55(1), pp. 6–35. doi:10.1177/0018720812457912.

Hudson, J. A., Zehner, G. F. & Meindl, R. S. (1998). "The USAF multivariate accommodation method." *SAE Technical Paper* 982833. SAE International. doi:10.4271/982833.

Human Factors and Ergonomics Society (2004). *Guidelines for using anthropometric data in product design*. Santa Monica, CA: HFES.

Human Factors and Ergonomics Society (2007). *ANSI/HFES 100–2007: Human Factors Engineering of Computer Workstations*. Santa Monica, CA: HFES.

Human Factors and Ergonomics Society (2020). *Multivariate Anthropometric Tool (MAT)* [software]. Available at: https://mat.hfes.org/.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-T., Rocktäschel, T., Riedel, S. & Kiela, D. (2020). "Retrieval-augmented generation for knowledge-intensive NLP tasks." In: *Advances in Neural Information Processing Systems* (NeurIPS 2020), vol. 33, pp. 9459–9474. Curran Associates, Inc. arXiv:2005.11401.

OpenAI (2023). *ChatGPT (June 26 version)*. Available at: https://chat.openai.com/chat.

OpenAI (2023). *GPT-4 [large language model]*. Available at: https://openai.com/gpt-4.

Parkinson, M. B. & Reed, M. P. (2010). "Creating virtual user populations by analysis of anthropometric data." *International Journal of Industrial Ergonomics*, 40, pp. 106–111. doi:10.1016/j.ergon.2009.07.003.

Petrat, D. (2021). "Artificial intelligence in human factors and ergonomics: an overview of the current state of research." *Discover Artificial Intelligence*, 1(3). doi:10.1007/s44163–021-00001–5.

Pheasant, S. & Haslegrave, C. M. (2006). *Bodyspace: Anthropometry, ergonomics and the design of work*. 3rd ed. Boca Raton, FL: Taylor & Francis.

Porter, J. M., Case, K., Marshall, R., Gyi, D. E. & Oliver, R. S. N. (2004). "'Beyond Jack and Jill': Designing for individuals using HADRIAN." *International Journal of Industrial Ergonomics*, 33, pp. 249–264. doi:10.1016/j.ergon.2003.08.002.

Robinette, K. M. & McConville, J. T. (1981). "An alternative to percentile models." *SAE Technical Paper* 810217. SAE International. doi:10.4271/810217.

Streamlit Inc. (2025). *Streamlit: The fastest way to build and share data apps*. Available at: https://streamlit.io/.

Sydor, M. & Hitka, M. (2023). "Chair size design based on user height." *Biomimetics*, 8(1). doi:10.3390/biomimetics8010057.

Van Rossum, G. & Drake, F. L. Jr. (1995). *Python reference manual*. Amsterdam: Centrum voor Wiskunde en Informatica.

Vega, G. R., Colado, U. Z., Colado, X. P. Z., Vega, D. A. R. & de la Vega Bustillos, E. J. (2021). "Comparison of univariate and multivariate anthropometric accommodation of the northwest Mexico population." *Ergonomics*, 64, pp. 1018–1034. doi:10.1080/00140139.2021.1892832.

Wang, M., Yu, B., Zhang, Y., Peng, C., Chen, Y., Zhou, W., Gu, J., Zhuang, C., Guo, R. & Zhao, X. (2025). "Function calling in large language models: Industrial practices, challenges, and future directions." *White paper*. Available at: https://github.com/Applied-Machine-Learning-Lab/Awesome-Function-Callings.