Making Bodies: Assumptions in the Design and Validation of Motion Capture Technology

Mona Sloane[1], **Abigail Z. Jacobs**[2], **Emanuel Moss**[1, 3]
[1]University of Virginia
[2]University of Michigan
[3]Intel Labs
mona.sloane@virginia.edu, azjacobs@umich.edu, emanuel.moss@intel.com

Abstract

Motion capture technologies are increasingly being incorporated into key aspects of social life, but embed potentially harmful assumptions. We examine the structural factors that allow for assumptions to solidify in motion capture systems and propose a matrix that works as a structured guide for motion capture developers to identify and document the assumptions baked into the system they are working on. To motivate novel conversations and future work in this area, we conclude with discussing the application of this matrix as well as its limitations.

1 Introduction

Recent years have been marked by heightened activity in AI regulation. The European Union, after years of development and deliberation, passed the EU AI Act which mandates risk-based assessments and mitigation strategies of AI systems across sectors (Edwards 2021; Meltzer and Tielemans 2022). Similarly, President Biden introduced the Executive Order on Safe, Secure and Trustworthy AI (Federal Register 2023) which sets out to regulate AI use across the US government. Despite not being a regulatory framework, the Executive Order had sweeping signaling effects for the tech industry. In 2025, the Biden Administration's Executive Order was replaced by the Trump Administration's Executive Order Removing Barriers to American Leadership in Artificial Intelligence, overturning specific policy directives but maintaining a policy interest in artificial intelligence for the US government. Despite recent, dramatic shifts in administration priorities, US agencies have long focused on mitigating potential AI harms. For example, the US Equal Employment Opportunity Commission (EEOC) has warned of potential disability discrimination in AI systems used in hiring and employment (Equal Employment Opportunity Commission 2-24) and the Federal Trade Commission (FTC) has strengthened its abilities in AI-related investigations of potential violations of consumer protection laws (Federal Trade Commission 2023a).

The US-based drugstore chain Rite Aid Corporation recently was prohibited by the Federal Trade Commision (FTC) from using facial recognition technology for five years (Federal Trade Commission 2023b). The FTC complaint showed that Rite Aid failed to put in place safeguards around facial recognition technology to protect store customers from harms, including false accusations, humiliations, and putting consumers' sensitive information at risk. The 2023 actors and writers strike in Hollywood evolved in large part around the use of AI systems in the film industry, including questions of ownership of motion capture and other biometric data of actors (Covey 2023). Both of these landmark cases involved the representation of human bodies as key components of AI systems, which points toward the need for careful attention to how motion capture technology, gesture recognition systems, and facial recognition products are developed.

Despite these regulatory developments and a heightened focus in the critical tech field on issues around fairness, accountability, and transparency in AI, there has been little attention within the motion capture space[1] on potential AI harms and AI risk mitigation (Pittaluga and Koppal 2016, is an exception). The ethical stakes of biometric data, in general, have been covered rather broadly (Awumey, Das, and Forlizzi 2024; Burema et al. 2023). However, the comparative lack of attention paid to motion capture reveals underexplored risks and harms that arise in other types of AI systems that make assumptions about the human body and its movement, resulting in motion capture design processes and practices that are not sufficiently engaged in risk and harm identification and mitigation.

It is evident that there is an urgent need to better understand, identify, and mitigate the potential harms of motion capture systems as they are rapidly being innovated and that, in doing so, lessons can be learned for AI systems more generally. Motion capture developers must develop a framework for exerting reasonable steps to ensure motion capture products are safe for use. This is particularly true for the AI harms that are the result of normative assumptions

[1]Here, we refer to *motion capture* as the set of technologies that are used to infer bodies and their movement from some kind of sensor (e.g., video, infrared camera systems detecting markers attached to the body, accelerometers worn on the body, infrared cameras external to the body). Motion capture technologies more broadly include tasks like activity, pose, or gesture recognition, that directly or indirectly depend on stylized models of bodies and their measurements.

around bodies and motion. Assumptions—for example, that a human body has four limbs and moves without aides—are baked into motion capture technology deeply and pervasively. Even for motion capture software that offers affordances for capturing motion of bodies with limb differences, other parts of the motion capture system infer measurements based on normative body types, in ways that are also central to the practices of developing and validating motion capture systems (Harvey et al. 2024). Typically, such assumptions go unscrutinized (Sloane, Moss, and Chowdhury 2022; Rhea et al. 2022; Barocas, Selbst, and Raghavan 2020; Jacobs and Wallach 2021). They become baked into the software, hardware, and training data and, importantly, benchmarks that make up motion capture systems.

Consequently, we propose that creating more fair, accountable, and transparent motion capture systems—and AI systems, generally—must be based on the systematic identification of the (potentially harmful) assumptions that underpin such systems and practices in design and validation. Such assumptions—and assumptions about representativeness—are the result of inherently social practices of measurement and AI development (Sloane and Zakrzewski 2022; Suchman et al. 2017) that shape what is measured, how data is chosen, and what features of data are relevant for making statistical inferences about a given phenomena. These choices, in turn, shape how data-driven systems might work better for some than others, or fail in certain contexts of use that are not adequately anticipated by designers. Here, sociotechnical approaches that substantively engage with validity of those assumptions can offer a path forward, especially for incorporating well-developed social scientific and mixed qualitative-quantitative methods into an analysis of risks and harms.

Below, we propose a systematic approach for tracing the hidden assumptions baked into motion capture systems. We first describe the structural factors that allow for assumptions to solidify in AI systems, and motion capture systems specifically, and consider their risks and potential harms. We then consider the implications for AI development and propose the motion capture matrix (Fig.1) as a structured guide for motion capture designers and innovators to identify and document the assumptions baked into the system they are working on. Such a matrix offers a tool for designers to both reflect on and document their assumptions, supporting critical analysis of motion capture systems and their technological descendants. The paper concludes with a discussion of future work needed to better identify and mitigate such motion capture technology risks moving forward, particularly in the design and validation phase.

2 Hidden Assumptions in Motion Capture Systems

Motion capture tasks, such as human motion prediction, pose estimation and activity recognition, create and depend on representations of human bodies. As a result, assumptions about what constitutes a human body—and whose bodies are typical or sufficiently representative—are inherently a part of motion capture design and validation. Potentially problematic assumptions about human bodies—that all humans are bilaterally symmetrical or that body proportions scale across heights and weights—can potentially produce representational harms when used in computer animation or documentary applications and material harms when used in medical or ergonomic applications.

Motion capture technology rests on the practice of measuring human bodies and inferring body shape and movement. Beginning in the 1930s, anthropometry, i.e., the measurement of bodies, was concerned with estimating body segment parameters to infer the force required to generate an observed motion (Harvey et al. 2024). Foundational work was funded by the US Air Force and sought to specify optimal cockpit design based on motion range, using a population less representative of live human motion—i.e., eight older, white male cadavers (Dempster 1955); a similar study used six cadavers(Chandler et al. 1975). Such work rests on crucial assumptions about the rigidity of human body segments that do not reflect reality. Yet, the measurements presented in both studies came to be accepted as "ground truth" for body segment parameter estimation in motion capture (Harvey et al. 2024). Harvey et al. (2024) have shown how subsequent research and innovation in motion capture used this "ground truth" to develop newly proposed formulas for body segment parameter estimation: The original assumptions about bodies and movement based on these very small and extremely biased samples became solidified in the hard- and software that makes up motion capture technology.

Today, the historical assumptions that underpin motion capture—specifically normative assumptions about body composition, proportions, and mobility (Campbell 2009)—remain relevant even for applications that do not directly make use of 3D motion capture technologies. This is because even 2D and depth camera approaches to motion capture are often developmentally linked to, trained on, and validated by earlier motion capture technologies. For instance, Human3.6M is commonly used for training and evaluation of 2D pose estimation tasks, and it relies on 3D marker-based motion capture data to parameterize joint positions and kinematic representation in their dataset, which themselves are derived through the motion capture apparatus manufacturer Vicon's proprietary "skeleton fitting procedure" (Ionescu et al. 2014).

This skeleton fitting procedure was developed by operationalizing the assumptions discussed above (Section 2), and therefore any applications developed using Human3.6M inherit these assumptions, even if they do not directly utilize Vicon hardware. In this way, the assumptions that historically underpin motion capture technology continue to shape the entire domain of motion, pose, and gesture recognition. Indeed, dataset developments such as Human3.6M also layer additional assumptions on top of earlier ones. The human actors who perform the gestures in this dataset represent a range of body mass indexes (BMI) that the authors assume captures "a moderate amount of body shape variability as well as different ranges of mobility" (Ionescu et al. 2014, p. 3). This assumption, that BMI differences correspond to body shape and mobility, however, is precisely the kind of assumption that calls for inspection as part of a process to ensure risks do not accrue according to body shape or mobility range upon deployment.

Datasets like Human3.6M or HumanEva-I, and the models built on top of them, address the challenge of inferring 2D representations of 3D poses from 2D data. However, having addressed core challenges of inferring such 2D skeletons and joint models, we now see the AI and motion capture community needs to bridge back to richer 3D representations of body shapes and positions (Elias, Sedmidubsky, and Zezula 2021; Lassner et al. 2017). Richer 3D representations of bodies are crucially needed as motion capture is used in applications such as sports and home and workplace safety. In closing this loop between 2D back to 3D, we see many of the same datasets serve as training and test in various aspects of the pipeline. The framing of the problems they solve (body shape estimation, gesture detection) and their evaluations are systematized based on past work (Zheng et al. 2023, for a current overview of 3D methods). For instance, much of this work on body shape and activity estimation infers 3D body shape using the seminal SMPL model ("A Skinned Multi-Person Linear Model") (Loper et al. 2015) or the related SMPLify method (Bogo et al. 2016). These papers, and the follow-on systems built open them, are then validated on datasets such as HumanEva-I and Human3.6M. Even additional 3D datasets include hardly more diverse populations (MPI-INF-3DHP is based on eight subjects (Mehta et al. 2017), and other datasets reuse those eight in various composite or synthetic datasets (Mehta et al. 2018, e.g.)). In addition, performance of new systems are also commonly then compared on their performance on Human3.6M (Martinez et al. 2017), thus carrying forward the assumptions and limitations built into Human3.6M, both in the design of the system and in the conceptualization of performance of the system.

Returning to the locus of hidden assumptions, consider how assumptions become embedded and are carried forward via the inference task, data collection, and validation in the following example. The SURREAL dataset (Varol et al. 2017), for instance, generated synthetic 3D and 2D data from other datasets including SMPL, poses from the CMU MoCap database (CMU Graphics Lab 2025), and body shapes from the 2002 CAESAR anthropometric survey (Robinette et al. 2002). The CMU MoCap database has no readily available demographic information; while the core database has 144 'subjects,' this corresponds to uniquely trained skeletons, not unique individuals, who were observed in a lab setting (CMU Graphics Lab 2025). In contrast, the CAESAR anthropometric survey, the data collection strategy—and the diversity of body shapes within—is articulated in their original report (Robinette et al. 2002):

> "The civilian populations of three countries were sampled in an effort to characterize the population of NATO countries as a whole. The United States was chosen because it has the largest and the most diverse population in NATO. The Netherlands was chosen because it has the tallest population in NATO, and Italy was chosen because it has one of the shortest populations in NATO."

As this quote makes clear, strong assumptions about representativeness of their population (including who is represented, what typical distributions are, and how they ought to be selected) are embedded in the CAESAR dataset, and therefore also any system that uses it anywhere in its pipeline, directly or indirectly.[2]

Here, assumptions around representativeness are implied as the statistical traits of a select sub-population are made to stand for a more diverse general population. Taken out of the context of CAESAR, such practices may introduce errors through dataset collection, where few bodies either stand in for the whole or are taken to define typical variation (Chasalow and Levy 2021; Wiley 2023; Scheuerman, Hanna, and Denton 2021). Used in such a way, the statistical models thus generated may be more error-prone for non-male, older, younger, individuals or people with disabilities. This in turn introduces significant risks where motion capture is deployed in safety-critical or health care scenarios like ergonomic design or kinesiology. Motion capture systems thus embody the same risks as other systems that essentialize social constructs (Scheuerman, Pape, and Hanna 2021), particularly over-generalizations about the typicality of bodies. This gives rise to new risks: in motion capture technologies meaningful differences between bodies are collapsed within a universal, 'essentially human' representation. This can, for example, become an issue in sports medicine, where biomechanical differences between adolescent and adult athletes could require specific statistical models for each population (Vaeyens et al. 2008). In such a scenario, adolescent athletes could be harmed or might not benefit as much as system developers had claimed, due to a fundamental epistemic mismatch between how their body is represented computationally and its actual position and movement through space.

3 Implications for AI Development

Potential harms in motion capture are magnified when one system is used as a benchmark to validate another system in the AI development process. Benchmarks have a diversity of meanings (Spendolini 1994) and of roles in the AI ecosystem (Paullada et al. 2021; Dotan and Milli 2020), and act a a site of assumptions becoming infrastructural (Denton et al. 2020) or a site for change (Bommasani 2023). Understanding the embedded assumptions is critically important for the key themes of "measurement, comparison, identification of best practices, implementation, and improvement" (Anand and Kodali 2008) that comprise any benchmarking and validation process.

In motion capture and related computer vision tasks, validation practices focus on various performance metrics such as accuracy, local similarity or identification of joints, and minimizing distances in comparisons to existing datasets, using datasets like Human3.6M as a benchmark (Zheng

[2]Of course, the CAESAR and CMU MoCap datasets are used widely beyond SURREAL, and SURREAL is not unique or exceptional in building off of these types of assumptions. Particularly as scholars look to synthetic data to as a workaround for otherwise more-homogeneous data training sets (Hu et al. 2022), these concerns follow.

et al. 2023, e.g., for an overview). The core practice here is that of *concurrent validity*, where new systems are compared to performance of previous iterations (Harvey et al. 2024). An emphasis on concurrent validity provides continuity and natural comparisons across systems and tasks; however, by design it also can overlook the ways that potentially undesirable or unintended assumptions that can ultimately impact real-world use. Systems that try to minimize distances to joint locations, for example, are perhaps generalizing—but only to humans that are similar to (or are captured similarly to) those in Human3.6M. Diversity of body types—as well as related attributes like age (particularly including children and the elderly), size, weight distribution, number of limbs, clothing type, accessories, activities, contexts, movement styles, hair styles, movement capabilities, and so on—made even more invisible, because systems that are thus labeled 'accurate' or 'validated' did not and can not capture such diversity—but are considered 'accurate' nonetheless.

Recall the SURREAL example in which the inference task centered on human body part segmentation and human depth estimation, comprising 3D tasks via inference from real and synthetic data (Varol et al. 2017). This task relied on assumptions about what makes human scenes realistic and what training data is needed, but (following good practice) they also considered validation beyond concurrent validity. In doing so, they revealed where hidden assumptions may have had an impact. For instance, they found that reliance on specific training sets yielded different outcomes, and that performance varied qualitatively across normative cultural and gendered clothes (Varol et al. 2017). While some assumptions remain unremarked upon or under-explored (e.g., body type from CAESAR) in their work, this type of qualitative exploration reveals that assumptions do matter, across task inference, data collection, and validation, and in ways that could map into real-world impacts.

Even though rarely made explicit in the design process, assumptions are traceable, as work from adjacent domains demonstrates (Moss 2021; Denton et al. 2021, 2020; Stevens and Keyes 2021), underlining that tracing and interrogating assumptions is important regardless of application domain. They present varying degrees of risk depending on the application domain, but may manifest as particularly high degrees of *harm* in domains that make assumptions about human bodies and behavior. Here, real-world harms can range from misrepresentation and distortions in entertainment-related applications to misdiagnosis in athletics or the medical field.

As regulation on AI is being mobilized to address potential AI harms, and motion capture becomes more pervasive across different domains (from surveillance to health monitoring, e.g.), a tractable approach is needed to help developers identify the normative assumptions baked into motion capture systems. We argue that such an approach should be *made integral to the practice of motion capture innovation* (Sloane and Moss 2022), i.e., be made integral to the already existing and overlapping ways of *doing* motion capture innovation and design, rather than become a box-ticking exercise that is imposed. The matrix can also, it should be noted, be used by those procuring motion capture systems or otherwise tasked with assessing their effectiveness within specific contexts of use (Sloane et al. 2021). To these ends, we propose an approach for responsible motion capture innovation that is inspired by social practice theory (Shove, Pantzar, and Watson 2012). Social practice theory examines how meanings, materials, and competencies solidify as the practices through which individuals coordinate their activities and interact with the world. Our approach centers on identifying the social practices of motion capture, tracing back assumptions about bodies and their movement, and considering the appropriateness of established benchmarks and validation practices throughout the AI development pipeline.

Focus	**Key Question/s**
Goal (E.g., diagnostics, clinical therapy, computer animation, gesture recognition, surveillance)	• What is the goal of the innovation (problem, practice, context)? • Why is motion capture technology the most suitable and appropriate approach for achieving this goal?
Motion Capture Approach (E.g., marker-based infrared camera systems, markerless video-based systems, or motion capture approaches based on inertial measurement units)	• What is the motion capture approach used? • How does the motion capture approach work and what is it optimizing for? • What parameters and choices are relevant to achieving that functionality?
Motion Capture Stage (Room Calibration, Subject Calibration, 2D Marker Identification, 3D Marker Reconstruction, Skeleton Solving)	• What motion capture stage/s are being innovated for? • What objects or outputs are produced in that stage that are passed onto the next stage? • What hardware and softwares are being built on?
Capture, Measurement, and Inference (E.g., body segments, skeleton positions, foot-strike and toe-off events)	• What is the capture focus? • What is the inference focus? • What are the measurement units and techniques? • What inference or classification tasks are being done?
Data (E.g., solved skeleton and markerset bones, rotation type, and movement classification, data format sampling frequency, data cleaning procedures, interpolation settings)	• What data, and what types of data, are used in training, testing, and operating the system? • How was the data processed and used? • How was the data collected and under what conditions? • What ethical protocols were adhered to?
Validation (E.g., assessment of accuracy on body error measurement, posture detection, gait analysis)	• What validation practices are used (for general system validation and for task validation)? • What validation practices were used to establish the benchmarks used? • What ground truth/s are used and how are they established?
Assumptions (E.g., Bodies are symmetrical, bodies are stable over time, body measurement errors are similar across body types)	• What are the assumptions baked into the system? • How and when did the assumptions stabilize? • What are the epistemological roots of the assumptions?

Table 1: The Motion Capture Matrix

We build on Sloane et al. (Sloane, Moss, and Chowdhury 2022) to propose a motion capture matrix that helps developers tease out the potentially problematic assumptions that are baked into motion capture technology. The matrix (Table 1) is broken down into the following focal points: goal; motion capture approach; motion capture stage; capture, measurements, and inference; data; validation; and assumptions. For each focus, the matrix proposes key questions that developers should ask in order to obtain and document the relevant information.

Development teams may use standard techniques for ethnographic data collection (Wasson 2000; Ladner 2016; Nardi 1997) in addition to literature research (and in some cases archival research) as well as consult any documentation that can be found on the specific motion capture systems, or components thereof, that they are using. They may also build on existing work that critically examines "ground truthing" (Moss 2021; Kang 2023), as well as established techniques for algorithmic impact assessments (Reisman et al. 2018; Moss et al. 2021) and tracing historical assumptions (Denton et al. 2020; Harvey et al. 2024; Scheuerman, Hanna, and Denton 2021; Shilton et al. 2021).

3.1 Goal

Often, technologies are innovated in the lab without clearly defined relationships to real-world problems, contexts, and conditions. This is particularly the case for AI-driven systems which still demonstrate a social (science) deficit (Sloane and Moss 2019). This persistent structural shortcoming routinely leads to detachments between technology development and context and social use (and usefulness). It is, in part, fuelled by stubborn expectation that technological approaches or AI techniques work well and are suitable for all kinds of different social contexts and use cases, rather than specific ones (Selbst et al. 2019). This can lead to harms and issues, ranging from systems that optimize for the wrong things (Barocas, Hardt, and Narayanan 2023) to systems that are accurate but biased (Obermeyer et al. 2019). To mitigate this dynamic, the first step in using the matrix is to clearly describe the goal of the (motion capture, in this case) innovation that is to be innovated or developed. This could be supported in diagnostics or clinical therapy, computer animation for entertainment, gesture recognition, or surveillance. Developers should clearly articulate what goal they are pursuing, i.e., what problem, practice, and context they are designing for. Here, they should rationalize why motion capture technology is the most suitable and appropriate approach.

3.2 Motion Capture Approach

The second step is to document and describe the particular motion capture approach that is to be developed. This is important as an early step as it defines the technological field in which the motion capture innovation is occurring, and it allows to trace how this approach was established and validated. The field of motion capture is rapidly evolving, and the list of potential motion capture types of technologies and approaches constantly grows. However, dominant types are marker-based infrared camera systems, markerless video-based systems, or motion capture approaches based on inertial measurement units (IMUs) which are devices that measure and report the specific gravity and angular rate of an object to which they are attached (Malleson et al. 2017). Included in the description of the motion capture approach should be system functionality, such as the use of infrared cameras for capturing 3D positions of markers to infer skeletons of bodies measured in an infrared capture studio, or algorithms assigning joint positions to 2D imagery. Here, developers should also rationalize their choice of parameters. For example, the choice of marker sets or skeletal segment mode, the frame rate, and the calibration tolerance.

3.3 Motion Capture Stage

Like many computational systems based on sensing, motion capture is modular. This modularity is reflected by an often hidden pipeline between data collection (e.g., from infrared signals or video), calibration (possibly not obvious to a system user), and the inference/analysis pipeline (e.g., detecting human body parts, and their simulation). Motion capture is modular in that motion capture systems are comprised of various stages and for each stage, a different approach or product/s may be used or developed. For instance, the marker-based motion capture systems often involve the following stages: room calibration, subject calibration, 2D marker identification, 3D marker reconstruction, and skeleton solving (Kirk, O'Brien, and Forsyth 2004; Topley and Richards 2020; Vicon 2022). Developers should describe for which stage they are innovating and what objects or outputs they are producing that are passed onto a subsequent stage (such as 2D marker identification that is reconstructed as 3D point locations in the next stage). If they are innovating across stages or are changing the motion capture workflow and its stages, they should document and rationalize that too. Additionally, they should also document which existing hardware and software they are building their innovation on, and which hardware and software previous and subsequent motion capture stages are likely to use or be built on.

3.4 Capture, Measurements, and Inference

Within each motion capture stage, there is a capture (what data is collected, which may go through some analysis pipeline) and downstream inference focus (what is inferred in order classify the desired object or phenomenon, such as bodies, movements, gestures, activities, or identity). All of these depend on a limited set of stylized body measurements that that are the basis for downstream inference – whether that inference is "healthy gait" (health status) or "waving hello" (activity) or demographic or identity prediction. Therefore, developers should clearly describe what the capture focus and inference focus is, and what measurements are at play. These could, for example, be body segments that are the basis for skeleton inference, skeleton positions that are the basis for pose estimation (Kirk, O'Brien, and Forsyth 2004; Zheng et al. 2023), or foot-strike and toe-off events are the basis for inferring ground reaction force (Patoz et al. 2021). Included in this description should be a documentation of the type of data collected and the techniques to capture them, the measurement units and techniques used, as well as the inference or classification tasks that are being done. This is particularly important since measurement practices solidify assumptions in motion capture technology (Harvey et al. 2024).

3.5 Data

Motion capture technologies are data-centric systems that use a wide array of data for training, testing and operating the system. In motion capture innovation, developers should describe the types of data that were used and how (including data cleaning protocols). Reuse of motion capture data may comprise solved skeletons and movement classification; data-relevant processes include data format, sampling frequency, data cleaning procedures, and interpolation settings. When developers use existing data, it is crucial for them to examine how that data was collected, i.e., whether standard norms of human subjects research apply and have been adhered to. When new data is collected, the same concerns should dictate the data collection process. Here, well-established frameworks in the field of data science ethics, notably Shilton et al. (Shilton et al. 2021), can point the way toward particularly suitable approaches to ensure appropriate collection, quality and diversity, and use of training data. Related efforts examining best practices around training data for computer vision tasks (Denton et al. 2020; Paullada et al. 2021; Scheuerman, Hanna, and Denton 2021; Scheuerman et al. 2023), a number of which are related to human body shape, pose, and motion inference related tasks, should inform appropriate data collection and use.

3.6 Validation

Validation practices are key to innovation in motion capture technology. They help to continually assess performance across a number of metrics. Similar to motion capture being a modular system, there also are different types of validation practices at play in designing motion capture systems. Like measurement practices, validation practices can further ingrain potentially problematic assumptions in motion capture systems, chiefly by way of ingraining those assumptions as ground truths against which newer systems are evaluated or benchmarked against (Harvey et al. 2024). Therefore, it is important that developers articulate what validation practices were used to justify the benchmarks they are using for their motion capture innovation, as well as describe clearly what validation practices they are using themselves. They should be mindful that multiple meaningful, complementary types of validation are needed in addition to the most common, concurrent validity, to consider the appropriateness of established benchmarks and validation practices writ large (Jacobs and Wallach 2021). In addition to concurrent validity, external validity should also be established, which requires going beyond performance on test data sets to consider how a system will be used in practice (e.g., from the lab setting to the field, from healthy actors to real users in diagnostic settings, or from university settings to low-resourced contexts (Harvey et al. 2024, for some examples)). Additionally, predictive validity (Jacobs and Wallach 2021) is important to show that systems produce outcomes related to the task of interest and can prevent harmful breakdowns. When using the matrix to be specific about the array of validation techniques they are using, developers should document what “ground truth” is used during validation, and how it is established.

3.7 Assumptions

Through the steps outlined above, developers can identify and articulate the assumptions that influence, or have influenced, data collection, system design, and validation practices. It is important to note that assumptions have epistemological roots. For example, many systems are based on stubborn assumptions about the suitability of proxies for the measurement of the desired construct, such as using joint angles as a proxy for fatigue in the development of ergonomic scores (Massiris Fernández et al. 2020). Often, it is possible to trace the roots of assumptions to certain eras and disciplines. For example, the idea that facial expressions or vocalization indicate criminality (Fussell 2020) has its roots in eugenics (Wiggins and Jones 2023; Levendowski 2021). Assumptions that have stabilized in the measurement and validation practices that make up motion capture today include assumptions such as *bodies are symmetrical*, *bodies are stable over time*, or *body measurement errors are similar across body types* (Harvey et al. 2024). They have their roots in ergonomics, physiognomy, and interaction design. Tracing assumptions may require carefully analyzing the methods sections of academic papers published alongside datasets, tracking down and inspecting the labels used in associated datasets, and interrogating the assumptions that would have needed to have been in place to produce labeled datasets.

Once developers have obtained and documented the information required by the motion capture matrix, they will be well equipped to anticipate potential harms of their designs and be in a better position to engage in new compliance, such as completing algorithmic impact assessments (Metcalf et al. 2021). By critically examining their design components, tools, and processes, they will also have engaged in reflexivity, putting them in a better position to make more ethical decisions, such as creating more meaningful benchmarks. Lastly, completing the matrix will serve as important documentation of motion capture technology design that can have both legal and historical significance.

4 Future Work

In this paper, we have argued that equitable, responsible, epistemically-sound development of any technical system depends on engaging with the often-hidden assumptions upon which they rest. This is particularly important for systems that have multiple stages and that include hardware components that carry their own assumptions, such as motion capture technology. Whilst the space of AI regulation is growing rapidly, motion capture has evaded both focused regulatory attention (unless treated specifically as biometrics for surveillance) and scrutiny by sociotechnical researchers. Therefore, we suggest that future research should apply and expand the approach laid out here. Similarly, research into what more meaningful validation and evaluation techniques can be established is necessary, alongside building out approaches for excavating hidden assumptions in AI systems. Cues can be taken from past critiques of stylized, circular evaluation tasks in related fields that have pointed to how to incorporate sensible, meaningful evaluation practices (Hutchinson et al. 2022, 2021; Bommasani 2023).

Akin to “ethical” or “broader impacts” statements that sometimes are required in other venues, explicitly *articulating* and *validating* assumptions as reasonable and legitimate is also a necessary condition for the broader research community to be able to examine, audit, challenge, and mitigate them. Such needs are already being called for by draft legislation and executive orders governing AI systems that call for this kind of work. In the U.S., the National Institute of Standards and Technology has produced an AI risk management framework whose dimensions facilitate this kind of work (Tabassi 2023). We suggest future work take this framework as a point of departure when deepening work on AI assumptions and motion capture technology. In this context, we want to note that interrogating the assumptions of AI systems is not without financial costs, which funding agencies should support as basic research, but which can be complemented by private institutions who benefit most from innovation in this domain.

Even though we have put much emphasis on increasing levels of transparency and accountability in motion capture technology design, we did exclude an important aspect: stakeholder engagement. All responsible design practices extend beyond data collection and should consider impacted stakeholders. Stakeholder engagement should not just follow narrow definitions of "user" groups, but consider who (i.e., individuals and communities) are possibly impacted by the envisioned use of the motion capture technology that is being developed. In accordance with such assessments, stakeholder groups should be meaningfully engaged in the research and innovation. Consequently, future work on fairness, accountability, and transparency in motion capture technology should encourage developers to be mindful of the significance of lived experience (Ferrarello 2021) and social practices (Suchman et al. 2017) that form the backdrop and context of the innovation they seek to make. In that context, they may wish to seek guidance from the disability community who has long–and successfully–grappled with these issues and questions (Hamraie and Fritsch 2019) whilst (co-)producing groundbreaking innovations such as speech-to-text and voice recognition systems, the curb cut, or bendy straws (Hoyt 2024).[3] Ultimately, the assumptions that underpin an undertaking may be found to be entirely inappropriate upon further study. Rather than constituting a failure or roadblock, such discoveries can be understood as direction-markers toward more fruitful, and better-grounded areas for research and innovation.

Acknowledgments

This project was supported by a Notre Dame-IBM Tech Ethics Lab grant. The authors thank Hauke Sandhaus and Emma Harvey for research support and Todd Bryant, Luke DuBois, Kaustav Sarkar, and Harsh Palan for their valuable support.

[3]Anyone seeking guidance on this topic is best served by initially reaching out to disability advocacy groups for pointers on community groups that might be implicated in a specific project. In the U.S., the Americans with Disabilities Act National Network coordinates regional centers that offer guidance for ADA compliance (http://www.adata.org) and Consortium for Citizens with Disabilities maintains a thorough list of advocacy organizations for developers to potentially partner with (http:www.https://www.c-c-d.org/).

References

Anand, G.; and Kodali, R. 2008. Benchmarking the benchmarking models. *Benchmarking: An international journal*, 15(3): 257–291.

Awumey, E.; Das, S.; and Forlizzi, J. 2024. A systematic review of biometric monitoring in the workplace: Analyzing socio-technical harms in development, deployment and use. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 920–932.

Barocas, S.; Hardt, M.; and Narayanan, A. 2023. *Fairness and machine learning: limitations and opportunities*. Cambridge, Massachusetts: The MIT Press. ISBN 978-0-262-04861-3.

Barocas, S.; Selbst, A. D.; and Raghavan, M. 2020. The hidden assumptions behind counterfactual explanations and principal reasons. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, 80–89.

Bogo, F.; Kanazawa, A.; Lassner, C.; Gehler, P.; Romero, J.; and Black, M. J. 2016. Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part V 14*, 561–578. Springer.

Bommasani, R. 2023. Evaluation for Change. In Rogers, A.; Boyd-Graber, J.; and Okazaki, N., eds., *Findings of the Association for Computational Linguistics: ACL 2023*, 8227–8239. Toronto, Canada: Association for Computational Linguistics.

Burema, D.; Debowski-Weimann, N.; von Janowski, A.; Grabowski, J.; Maftei, M.; Jacobs, M.; van der Smagt, P.; and Benbouzid, D. 2023. A sector-based approach to AI ethics: Understanding ethical issues of AI-related incidents within their sectoral context. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '23, 705–714. New York, NY, USA: Association for Computing Machinery. ISBN 9798400702310.

Campbell, F. 2009. *Contours of ableism: The production of disability and abledness*. Springer.

Chandler, R.; Clauser, C.; McConville, J.; Reynolds, H.; and Young, J. 1975. Investigation of Inertial Properties of the Human Body. Technical Report ADA016485, Air Force Aerospace Medical Research Lab.

Chasalow, K.; and Levy, K. 2021. Representativeness in statistics, politics, and machine learning. In *Proc. ACM FAccT*, 77–89.

CMU Graphics Lab. 2025. CMU Graphics Lab Motion Capture Database. http://mocap.cs.cmu.edu/. Accessed:2025-05-21.

Covey, L. 2023. Hollywood strikes fight over ownership of biometrics. In *Cyber Protection Magazine*; accessed 2025-05-21.

Dempster, W. T. 1955. Space requirements of the seated operator: geometrical, kinematic, and mechanical aspects of the body, with special reference to the limbs. Technical Report WADC TECHNICAL REPORT 55-159, Wright Air Development Center.

Denton, E.; Hanna, A.; Amironesei, R.; Smart, A.; and Nicole, H. 2021. On the genealogy of machine learning datasets: A critical history of ImageNet. *Big Data & Society*, 8(2): 20539517211035955.

Denton, E.; Hanna, A.; Amironesei, R.; Smart, A.; Nicole, H.; and Scheuerman, M. K. 2020. Bringing the People Back In: Contesting Benchmark Machine Learning Datasets. ArXiv:2007.07399 [cs].

Dotan, R.; and Milli, S. 2020. Value-laden disciplinary shifts in machine learning. *Proc ACM Fairness Accountability and Transparency*.

Edwards, L. 2021. The EU AI Act: a summary of its significance and scope. *Artificial Intelligence (the EU AI Act)*, 1.

Elias, P.; Sedmidubsky, J.; and Zezula, P. 2021. Understanding the limits of 2D skeletons for action recognition. *Multimedia Sys.*, 27: 547–561.

Equal Employment Opportunity Commission. 2-24. Disability Discrimination. https://www.eeoc.gov/employers/small-business/disability-discrimination. Accessed 2024-01-25.

Federal Register. 2023. Executive Order 14110. *Federal Register*, 88(210): 75191–75226.

Federal Trade Commission. 2023a. FTC Authorizes Compulsory Process for AI-related Products and Services.

Federal Trade Commission. 2023b. Rite Aid Banned from Using AI Facial Recognition After FTC Says Retailer Deployed Technology without Reasonable Safeguards.

Ferrarello, S. 2021. *Phenomenology of bioethics: technoethics and lived-experience*, volume 84. Springer.

Fussell, S. 2020. An Algorithm That 'Predicts' Criminality Based on a Face Sparks a Furor. *Wired*. Section: tags.

Hamraie, A.; and Fritsch, K. 2019. Crip Technoscience Manifesto. *Catalyst: Feminism, Theory, Technoscience*, 5(1): 1–33. Number: 1.

Harvey, E.; Sandhaus, H.; Jacobs, A. Z.; Moss, E.; and Sloane, M. 2024. *The Cadaver in the Machine*: The Social Practices of Measurement and Validation in Motion Capture Technology. *Proc. ACM CHI*.

Hoyt, A. 2024. 8 Everyday Items Originally Invented for People With Disabilities. In *HowStuffWorks*; date accessed 2025-05-21; article last updated 2024-02-27.

Hu, H.; Xiao, D.; Rhodin, H.; and Murphy, T. H. 2022. Towards a Visualizable, De-identified Synthetic Biomarker of Human Movement Disorders. *Journal of Parkinson's Disease*, 12(7): 2085–2096.

Hutchinson, B.; Rostamzadeh, N.; Greer, C.; Heller, K.; and Prabhakaran, V. 2022. Evaluation gaps in machine learning practice. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 1859–1876.

Hutchinson, B.; Smart, A.; Hanna, A.; Denton, E.; Greer, C.; Kjartansson, O.; Barnes, P.; and Mitchell, M. 2021. Towards accountability for machine learning datasets: Practices from software engineering and infrastructure. In *Proc. ACM FAccT*, 560–575.

Ionescu, C.; Papava, D.; Olaru, V.; and Sminchisescu, C. 2014. Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7): 1325–1339.

Jacobs, A. Z.; and Wallach, H. 2021. Measurement and Fairness. In *Proc. ACM FAccT*, 375–385. ACM.

Kang, E. B. 2023. Ground truth tracings (GTT): On the epistemic limits of machine learning. *Big data & society*, 10(1): 20539517221146122.

Kirk, A.; O'Brien, J. F.; and Forsyth, D. A. 2004. Skeletal parameter estimation from optical motion capture data. In *ACM SIGGRAPH 2004 Sketches*, 29.

Ladner, S. 2016. *Practical ethnography: A guide to doing ethnography in the private sector*. Routledge.

Lassner, C.; Romero, J.; Kiefel, M.; Bogo, F.; Black, M. J.; and Gehler, P. V. 2017. Unite the people: Closing the loop between 3D and 2D human representations. In *Proc. IEEE CVPR*, 6050–6059.

Levendowski, A. 2021. Face Surveillance Was Always Flawed.

Loper, M.; Mahmood, N.; Romero, J.; Pons-Moll, G.; and Black, M. J. 2015. SMPL: a skinned multi-person linear model. *ACM Trans. Graph.*, 34(6).

Malleson, C.; Gilbert, A.; Trumble, M.; Collomosse, J.; Hilton, A.; and Volino, M. 2017. Real-time full-body motion capture from video and imus. In *2017 international conference on 3D vision (3DV)*, 449–457. IEEE.

Martinez, J.; Hossain, R.; Romero, J.; and Little, J. J. 2017. A Simple Yet Effective Baseline for 3D Human Pose Estimation. In *2017 IEEE International Conference on Computer Vision (ICCV)*, 2659–2668. Venice: IEEE. ISBN 978-1-5386-1032-9.

Massiris Fernández, M.; Fernández, J. Á.; Bajo, J. M.; and Delrieux, C. A. 2020. Ergonomic risk assessment based on computer vision and machine learning. *Computers & Industrial Engineering*, 149: 106816.

Mehta, D.; Rhodin, H.; Casas, D.; Fua, P.; Sotnychenko, O.; Xu, W.; and Theobalt, C. 2017. Monocular 3D human pose estimation in the wild using improved cnn supervision. In *Proc. 3D Vision (3DV)*, 506–516. IEEE.

Mehta, D.; Sotnychenko, O.; Mueller, F.; Xu, W.; Sridhar, S.; Pons-Moll, G.; and Theobalt, C. 2018. Single-Shot Multi-Person 3D Pose Estimation From Monocular RGB. In *Proc. 3D Vision (3DV)*. IEEE.

Meltzer, J.; and Tielemans, A. 2022. The European Union AI Act. *Bruselj: Brookings Institution*.

Metcalf, J.; Moss, E.; Watkins, E. A.; Singh, R.; and Elish, M. C. 2021. Algorithmic impact assessments and accountability: The co-construction of impacts. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, 735–746.

Moss, E.; Watkins, E. A.; Metcalf, J.; and Elish, M. C. 2021. Governing with algorithmic impact assessments: six observations. In *AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society (AIES)*.

Moss, E. D. 2021. *The Objective Function: Science and Society in the Age of Machine Intelligence*. Ph.D. thesis, City University of New York.

Nardi, B. A. 1997. The use of ethnographic methods in design and evaluation. In *Handbook of human-computer interaction*, 361–366. Elsevier.

Obermeyer, Z.; Powers, B.; Vogeli, C.; and Mullainathan, S. 2019. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464): 447–453.

Patoz, A.; Lussiana, T.; Gindre, C.; and Malatesta, D. 2021. A novel kinematic detection of foot-strike and toe-off events during noninstrumented treadmill running to estimate contact time. *Journal of Biomechanics*, 128: 110737.

Paullada, A.; Raji, I. D.; Bender, E. M.; Denton, E.; and Hanna, A. 2021. Data and its (dis)contents: A survey of dataset development and use in machine learning research. *Patterns*, 2(11): 100336.

Pittaluga, F.; and Koppal, S. J. 2016. Pre-capture privacy for small vision sensors. *IEEE transactions on pattern analysis and machine intelligence*, 39(11): 2215–2226.

Reisman, D.; Schultz, J.; Crawford, K.; and Whittaker, M. 2018. Algorithmic Impact Assessments: A Practical Framework for Public Agency. *AI Now*.

Rhea, A. K.; Markey, K.; D'Arinzo, L.; Schellmann, H.; Sloane, M.; Squires, P.; Kahn, F. A.; and Stoyanovich, J. 2022. An External Stability Audit Framework to Test the Validity of Personality Prediction in AI Hiring. ArXiv:2201.09151 [cs].

Robinette, K. M.; Blackwell, S.; Daanen, H.; Boehmer, M.; Fleming, S.; Brill, T.; Hoeferlin, D.; and Burnsides, D. 2002. Civilian American and European surface anthropometry resource (CAESAR), final report, volume I: Summary. *Sytronics Inc Dayton Oh*.

Scheuerman, M. K.; Hanna, A.; and Denton, E. 2021. Do datasets have politics? Disciplinary values in computer vision dataset development. *Proc. ACM CSCW*, 5(CSCW2): 1–37.

Scheuerman, M. K.; Pape, M.; and Hanna, A. 2021. Auto-essentialization: Gender in automated facial analysis as extended colonial project. *Big Data & Society*, 8(2): 20539517211053712.

Scheuerman, M. K.; Weathington, K.; Mugunthan, T.; Denton, E.; and Fiesler, C. 2023. From Human to Data to Dataset: Mapping the Traceability of Human Subjects in Computer Vision Datasets. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW1): 1–33.

Selbst, A. D.; Boyd, D.; Friedler, S. A.; Venkatasubramanian, S.; and Vertesi, J. 2019. Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency*, 59–68.

Shilton, K.; Moss, E.; Gilbert, S. A.; Bietz, M. J.; Fiesler, C.; Metcalf, J.; Vitak, J.; and Zimmer, M. 2021. Excavating awareness and power in data science: A manifesto for trustworthy pervasive data research. *Big Data & Society*, 8(2): 20539517211040759.

Shove, E.; Pantzar, M.; and Watson, M. 2012. *The dynamics of social practice: Everyday life and how it changes*. Sage.

Sloane, M.; Chowdhury, R.; Havens, J. C.; Lazovich, T.; and Rincon Alba, L. 2021. AI and procurement-a primer.

Sloane, M.; and Moss, E. 2019. AI's social sciences deficit. *Nature Machine Intelligence*, 1(8): 330–331.

Sloane, M.; and Moss, E. 2022. Introducing a Practice-Based Compliance Framework for Addressing New Regulatory Challenges in the AI Field. *TechReg Chronicle*, (March): 9.

Sloane, M.; Moss, E.; and Chowdhury, R. 2022. A Silicon Valley love triangle: Hiring algorithms, pseudo-science, and the quest for auditability. *Patterns*, 3(2): 100425.

Sloane, M.; and Zakrzewski, J. 2022. German AI Start-Ups and "AI Ethics": Using A Social Practice Lens for Assessing and Implementing Socio-Technical Innovation. In *2022 ACM Conference on Fairness, Accountability, and Transparency*, 935–947. Seoul Republic of Korea: ACM. ISBN 978-1-4503-9352-2.

Spendolini, M. 1994. *The Benchmarking Book*. EBSCO ebook academic collection. Amacom. ISBN 9780814478660.

Stevens, N.; and Keyes, O. 2021. Seeing infrastructure: Race, facial recognition and the politics of data. *Cultural Studies*, 35(4-5): 833–853.

Suchman, L.; Blomberg, J.; Orr, J. E.; and Trigg, R. 2017. Reconstructing technologies as social practice. In *The Anthropology of Organisations*, 431–447. Routledge.

Tabassi, E. 2023. AI Risk Management Framework: AI RMF (1.0). Technical Report NIST AI 100-1, National Institute of Standards and Technology, Gaithersburg, MD.

Topley, M.; and Richards, J. G. 2020. A comparison of currently available optoelectronic motion capture systems. *Journal of biomechanics*, 106: 109820.

Vaeyens, R.; Lenoir, M.; Williams, A. M.; and Philippaerts, R. M. 2008. Talent identification and development programmes in sport: current models and future directions. *Sports medicine*, 38: 703–714.

Varol, G.; Romero, J.; Martin, X.; Mahmood, N.; Black, M. J.; Laptev, I.; and Schmid, C. 2017. Learning from synthetic humans. In *Proc. IEEE CVPR*, 109–117.

Vicon. 2022. Motion capture buyer's guide for life sciences.

Wasson, C. 2000. Ethnography in the field of design. *Human organization*, 59(4): 377–388.

Wiggins, C.; and Jones, M. L. 2023. *How Data Happened: A History from the Age of Reason to the Age of Algorithms*. WW Norton & Company.

Wiley, A. S. 2023. Biological normalcy. *Annual Review of Anthropology*, 52: 223–238.

Zheng, C.; Wu, W.; Chen, C.; Yang, T.; Zhu, S.; Shen, J.; Kehtarnavaz, N.; and Shah, M. 2023. Deep learning-based human pose estimation: A survey. *ACM Computing Surveys*, 56(1): 1–37.