# Don't call it *privacy-preserving* or *human-centric* pose estimation if you don't measure privacy

**Michele Baldassini*** **Francesco Pistolesi*** **Beatrice Lazzerini**
Department of Information Engineering
University of Pisa
Largo Lazzarino, 1 – 56122 Pisa (IT)
michele.baldassini@ing.unipi.it, {francesco.pistolesi, beatrice.lazzerini}@unipi.it

## Abstract

This position paper argues that human pose estimation (HPE) cannot be considered privacy-preserving or human-centric unless privacy is measured and evaluated. Although privacy concerns have become more visible in recent years, HPE systems are still assessed almost exclusively using *accuracy* metrics. Privacy is neither defined in measurable terms nor linked to regulatory requirements, and common deployment architectures introduce additional risks due to data transmission and storage. We highlight the limitations of current practices, including the continued reliance on RGB inputs and the lack of benchmarks that reflect legal and ethical constraints. We call for a shift in evaluation practices: *privacy* must become part of how HPE systems are designed, tested, and compared.

## 1 Introduction

**Position: This paper argues that human pose estimation must shift from performance-only evaluation to human-centric approaches that measure and prioritize privacy.**

Human pose estimation (HPE) determines the position of human joints and connections from images or videos. It has achieved impressive results in the last decade, becoming key in computer vision with applications to human-computer interaction [1], surveillance [2], entertainment [3], sports analytics [4], and healthcare [5]. These advances have been facilitated by powerful deep learning architectures [6] and large-scale, richly annotated datasets containing RGB images or videos collected in ideal [7–9] or synthetic [10–12] settings. Recent years have seen growing concerns about privacy, as AI models can expose *sensitive personal information (SPI)*, for example, facial features, body morphology, gender, or ethnicity [13]. This led to exploring anonymization methods and data types that capture less sensitive information than RGB images, such as LiDAR scans, thermal images, and depth maps.

Real-world HPE applications, including workplace safety, ergonomic risk assessment, physical rehabilitation, and elderly care, require the deployment of HPE systems in environments where privacy, regulation, and practical constraints are mandatory. These domains involve detailed, continuous monitoring of human movement, and have promoted datasets such as *IKEA ASM* [14] (multi-view recordings of furniture assembly tasks), *UCO-Labeled* [15] (physical exercises), and *MMRI* [16] (multimodal rehabilitation data). These datasets assume that accurate pose estimation requires detailed visual input, often revealing SPI.

This assumption is problematic, as privacy protection is a legal, ethical, and social obligation in real-world contexts. In particular, regulations require data minimization, purpose limitation, and explicit consent, thus making the acquisition of sensitive image data incompatible with these principles.

*Equal contribution.

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Position Paper Track.

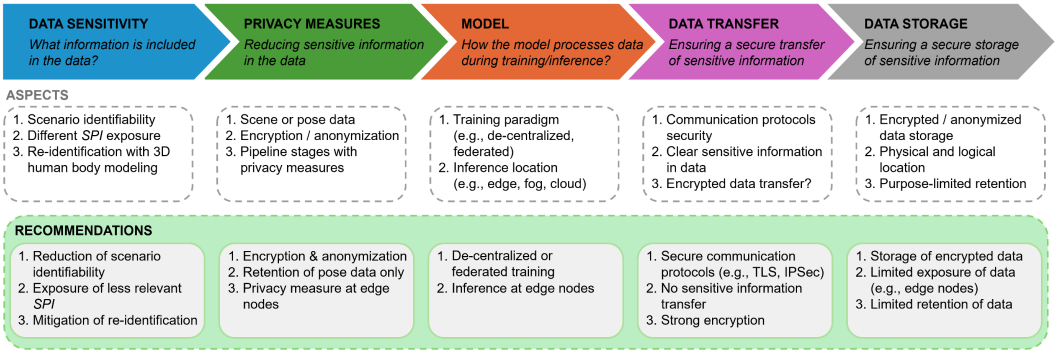


Figure 1: Overview of aspects and recommendations related to some key dimensions for privacy assessment.

Privacy concerns also involve the representation models of the body used in HPE. For example, the *kinematic model* only provides joint coordinates and is relatively privacy-preserving; the *planar model*, based on silhouettes or statistical shapes, reveals more about body posture and form; the *volumetric model*, used in 3D mesh reconstruction, encodes body shape and movement with enough fidelity to infer gender, age, or health conditions, thus posing serious privacy concerns.

Many modern HPE systems achieve high accuracy, but continue to use RGB data, which contain highly identifiable features [17–19]. This leads to serious privacy risks, including re-identification and unauthorized use. In addition, real-world HPE system deployments rely on *distributed architectures* that process data across multiple layers to distribute the computational load. Privacy risks are thus increased, as sensitive data are transferred across multiple layers—for example, from local acquisition and inference (edge) to centralized aggregation or training (cloud). Protecting sensitive data across system layers requires long-term privacy measures. HPE systems have started to address these challenges using techniques such as obfuscation, alternative data types (e.g., Wi-Fi, thermal, LiDAR), secure transmission and computation (e.g., TLS, encryption), and decentralized training (e.g., federated learning). These methods reduce privacy risks while supporting accuracy and legal compliance. Still, balancing privacy, performance, and deployment constraints remains a challenge.

Although privacy has received more attention in recent HPE systems, their performance is only measured using precision indicators for 2D HPE [20, 21] and 3D HPE [22–27]. Privacy is evaluated without using metrics or referring to regulatory standards, while also overlooking the risks due to data transit, storage, and processing when deploying HPE systems on multi-tier architectures, e.g., *edge-fog-cloud* [28]. Figure 1 shows, from left to right, some key dimensions that characterize many modern HPE systems, the aspects to consider for each dimension, and some recommendations to address to increase the level of privacy.

These limitations highlight a gap between current research practices and the need for privacy-aware, real-world applications. Addressing this gap requires defining privacy as a measurable aspect of HPE, and making it part of the system evaluation.

We invite the NeurIPS community to play an active role in redefining the foundations of HPE. We do not lack technical capabilities, but we lack standardized benchmarks, community incentives, and a shared vision grounded in real deployment contexts. We outline the limitations of current approaches, identifying key ethical and regulatory gaps. Finally, we reflect on how to design, develop, and deploy human-centric datasets and models that combine performance and privacy.

The paper is structured as follows: Section 2 explores alternative solutions to preserve privacy in HPE; Section 3 presents a view on possible ways for measuring the privacy risk when designing and deploying HPE systems; Section 4 discusses some alternative views; Section 5 concludes the paper.

## 2 How to preserve privacy?

The shift toward privacy-preserving modalities and system architectures often entails a trade-off in positional accuracy—particularly in the localization of fine-grained joint coordinates. Privacy-driven

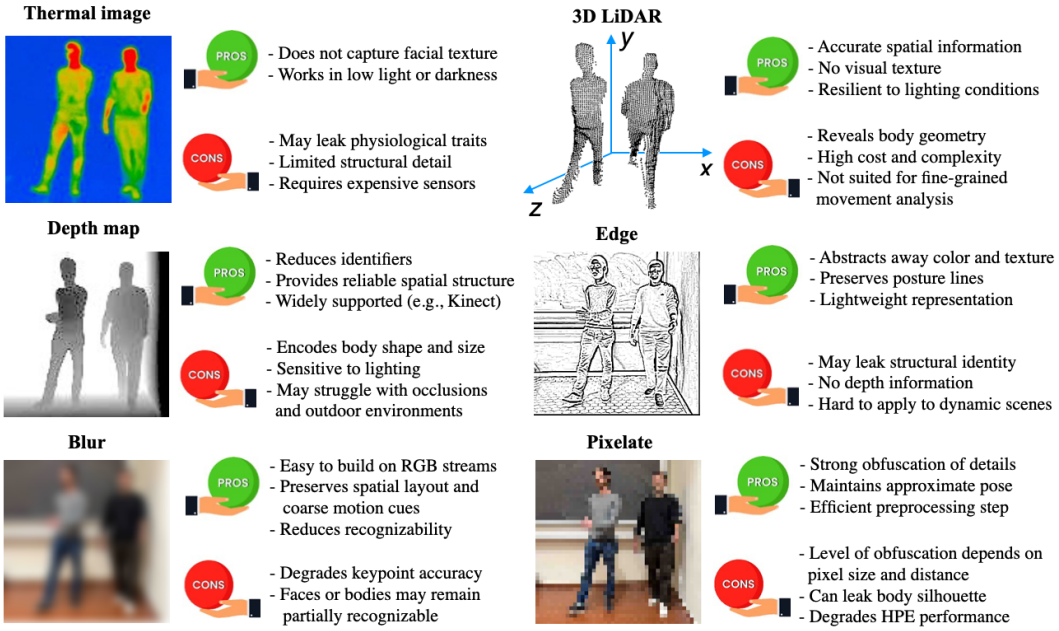


Figure 2: Examples of data modalities used in privacy-aware pose estimation, each with some pros and cons in terms of accuracy and privacy.

obfuscation, reduced resolution, and sensor substitution (e.g., depth in place of RGB) can significantly degrade the precision of joint-level estimations. This, in turn, affects downstream computations such as joint angle estimation, body segment alignment, and movement smoothness—metrics that are essential in domains governed by regulatory standards.

### 2.1 Changing modality

While RGB-based datasets dominate the field for their high accuracy, their inherent identifiability raises significant ethical and legal concerns. Conversely, alternative modalities (depth, LiDAR, thermal, RF) offer privacy-preserving properties but vary widely in usability, generalizability, and task-specific performance. This section argues that no single modality is universally "best" for privacy; instead, the field urgently requires a standardized measurement framework to evaluate trade-offs across modalities systematically.

Recent HPE datasets include privacy-preserving modalities—such as depth, LiDAR, thermal, and RF—often alongside RGB (see Figure 2). Although multimodal datasets such as NTU-RGBD [29,30] and LiDARHuman26M [31] show high performance with non-visual data, RGB datasets, e.g., COCO [11], continue to dominate due to their accuracy, despite high identifiability risks.

Synthetic datasets (e.g., SURREAL [32]) and RF/thermal datasets (e.g., MM-FI [33]) offer promising privacy advantages, yet lack scale, realism, or consistent evaluation. High-precision datasets like Human3.6M [34] provide detailed annotations but amplify privacy concerns.

Non-visual sensing modalities, such as radio frequency (RF), depth, and thermal imaging, are often presented as privacy-preserving alternatives to RGB. However, these technologies maintain sensitive personal information, and there is currently no standard metric to compare their privacy levels.

**Radio frequency (RF).** RF-based HPE methods—including RFID, FMCW radar, mmWave radar, and Wi-Fi—enable pose detection without visual data, providing anonymity [35–47]. While avoiding facial or direct visual biometric capture, these systems extract detailed skeletal and movement patterns that can act as unique biometric identifiers. This creates privacy risks related to covert tracking, re-identification, and behavioral profiling, especially in private or sensitive environments. Recent advances improve pose detail, increasing the sensitivity of collected biometric data and emphasizing the need for strong privacy safeguards.

Table 1: Modalities with associated pose representations and privacy risks.

| Modality | Pose Representation | Privacy Risks |
|---|---|---|
| RGB images | 2D skeletons<br>full-body meshes | High identifiability; re-identification via appearance cues, clothing, background context; face recognition leakage. Dominates accuracy but raises ethical/legal concerns [11, 29, 30]. |
| RF<br>Wi-Fi | 2D/3D skeletons<br>motion trajectories | Skeletal/motion info can reveal unique biometric signatures [94]; covert tracking [95]; re-identification; behavioral profiling [96]; environmental noise does not fully prevent re-identification [48–50]. |
| Depth sensing | 2D/3D skeletons<br>voxel grids<br>meshes | Body-shape re-identification; motion patterns and behavioral profiling; fusion with RGB data increases privacy risks [59–61, 73, 74]. |
| Thermal sensing | 2D/3D skeletons<br>meshes | Body-shape re-identification; motion patterns, behavioral profiling and physiological traits (e.g., emotion and temperature); fusion with other modalities increases risk [83–85, 88, 89]. |

**Wi-Fi.** Wi-Fi-based HPE uses reflections of standard Wi-Fi signals to infer body poses without visual cameras [48–58]. However, despite the environmental factors, signal noise, and lower spatial resolution compared to radar systems, Wi-Fi HPE can still capture detailed skeletal and motion information that reveal unique biometric patterns. Therefore, strong privacy safeguards are essential to prevent misuse of pose data and protect individuals' anonymity in Wi-Fi sensing applications.

**Depth sensing.** Depth map-based HPE reconstructs 3D body poses without RGB images [59–76], yet the detailed skeletal data it generates can serve as biometric identifiers, enabling person recognition and behavioral profiling. Fusion with RGB data further elevates privacy risks [73, 74].

LiDAR-based HPE provides high-resolution 3D body reconstructions [77–82], capturing uniquely identifiable body shapes and motions. Its growing use in the public and medical domains raises concerns about covert tracking and unauthorized surveillance.

Both modalities extract biometric signatures from non-visual data, necessitating strict privacy protections to prevent misuse and ensure anonymity.

**Thermal sensing.** Thermal imaging enables HPE without visible light, offering some privacy benefits [83]. However, it still reveals body shapes and motion patterns that can serve as biometric identifiers, risking re-identification. Additionally, thermal data expose physiological traits such as emotional states and environmental responses [84, 85], raising further privacy concerns.

Although lower in resolution, thermal images combined with depth data improve accuracy, but increase the risk of exposing identifiable features [86, 87]. Adaptations of visible-light pose models to thermal imaging [88–92] confirm persistent risks of biometric and physiological data leakage. Large annotated datasets [88, 89, 93] facilitate development but also highlight the need for careful privacy protection.

Thermal imaging partially protects visual identity, but still reveals sensitive biometric and physiological information, requiring strict privacy measures.

Table 1 summarizes the sensing modalities, their pose representations, and their privacy risks.

### 2.2 Anonymize visual data

Reducing the amount of identifiable information in visual data is an easy way to improve privacy in HPE systems. However, anonymization is not a guarantee. Each technique has trade-offs, and many preserve residual cues that can still identify individuals.

Lowering image resolution hides facial features and fine details, but often leaves body shape, gait, and motion patterns intact. These structural signals are enough to re-identify subjects in many cases, especially when combined with temporal information [97–99].

Software-based obfuscation techniques such as blurring or noise injection (for example, by using pixelation) are easy to apply and effective at masking sensitive visual content (see the bottom line of Figure 2). However, they degrade the quality of keypoint detection, thereby reducing accuracy. Also,


Figure 3: Comparison of *kinematic*, *planar*, and *volumetric* human pose representations (from left to right), showing the relation between abstraction level, identifiable information, and privacy risks.

obfuscated data are not safe by default: several studies show that neural models can reconstruct or infer masked information using priors or multi-frame context [100–102].

Hardware-based alternatives, such as event cameras, infrared sensors, or low-resolution depth cameras, limit the amount of biometric information captured at the source [103–105]. This can reduce the need for obfuscation, but has limitations including cost, and residual physiological traces—for example, heat signatures or breathing motion—that can still leak private information.

Hybrid setups exist that combine the strengths of both approaches. For example, some apply optical filters or hardware constraints during acquisition, followed by software processing [106, 107]. Although these hybrid techniques reduce direct exposure, they still require access to partially informative raw data at some stage in the pipeline, which reintroduces privacy risks.

Techniques like federated learning can help reduce data exposure by keeping raw data close to the acquisition point [108]. But again, this does not guarantee privacy: model updates can leak sensitive features, and attacks based on gradients or reconstruction from weights remain an open concern.

Thus, anonymization is not a solution. Systems that are claimed to preserve privacy should prove which types of personal information are still accessible after processing, and under what conditions.

### 2.3 Adopting privacy-aware pose representations

The way we represent human pose has important consequences for privacy. Different model types expose specific levels of personal information.

For example, kinematic models, often used in both 2D and 3D pose estimation, represent the body as a set of joint positions connected by limbs. These models focus on structure and motion, not appearance, making them relatively privacy-friendly. Planar models, which use silhouettes or simplified contours, provide more visual information about posture and body shape [109]. This can reveal physical traits that may enable identification or profiling. Volumetric models generate full-body 3D meshes with shape and motion details [110–112]. These representations can encode sensitive attributes such as gender, age, or physical condition, even without RGB input. Figure 3 shows the advantages and privacy risks of each model based on the level of abstraction and identifiable information retained.

Choosing a pose representation is thus both a technical and a privacy decision. We should consider what level of detail is needed for the task at hand, and whether a simpler representation could meet the goal while better protecting privacy.

Thus, pose estimation is privacy-preserving only if we can measure privacy and consider it at every stage: data lifecycle, model training and evaluation, and final deployment.

## 3 How can we start measuring privacy?

This section gives some directions for evaluating privacy in HPE. Starting from legal requirements, we give an example of how each part of an HPE system (data, model, inference, and transmission) contributes to the overall privacy risk.

### 3.1 Privacy risk factors

Privacy risk in HPE depends on various design choices, including:

- the type of *scene data* collected and how *human pose* is represented;

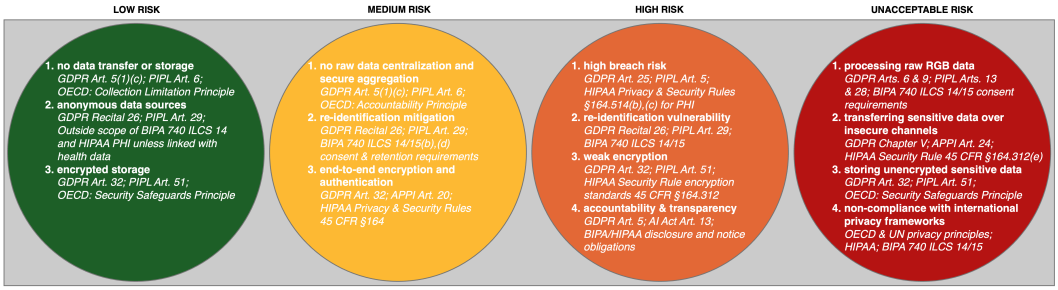


Figure 4: Risk levels in HPE systems and examples of privacy practices. Each circle represents a category of privacy exposure ranging from low (green) to unacceptable (red), with legal references from GDPR, PIPL, APPI, and BIPA.

- where and how *inference* is performed;
- where and how *training* is performed, and how model updates are distributed;
- where the *storage* takes place, which data are stored and for how long;
- how data *transmission* occurs and how many devices are involved in transfer and storage.

For example, a model trained on local data and never exposed to raw video is safer than one trained on images stored remotely. If personal data are kept longer than needed, or reused for other purposes, the risk is increased.

### 3.2 A risk-based view based on regulations

Legal frameworks such as the *General Data Protection Regulation (GDPR)* , the *AI Act*, the *Personal Information Protection Law (PIPL)*, and the *Biometric Information Privacy Act (BIPA)* part of the *Illinois Compiled Statutes (ILCS)* provide clear principles to handle personal and biometric data. These include data minimization, purpose limitation, storage limitation (GDPR Art. 5), privacy by design and by default (Art. 25), and restrictions on the use of biometric data (Art. 9). The AI Act defines high-risk systems (Annex III), including those used for biometric categorization or workplace monitoring.

Similar rules appear in China's PIPL (Art. 28 and 51) and Japan's *Act on the Protection of Personal Information (APPI)* (Art. 23 and 24), which also limit data retention and international transfers.

In the U.S., the BIPA imposes strict requirements on the collection and storage of biometric identifiers (ILCS chapter 740 Act. 14/15). Also, the *Health Insurance Portability and Accountability Act (HIPAA)* establishes standards for the protection, use, and disclosure of any identifiable health data (*protected health information (PHI)*) handled by covered entities and their associates.

At an international level, the *Organization for Economic Co-operation and Development (OECD)* and *United Nations (UN)* guidelines promote principles such as accountability, transparency, and proportionality in AI and data processing.

Due to their widespread adoption worldwide, we should use these principles as a starting point to evaluate the privacy risk of HPE systems in years to come.

We try to do this by outlining four levels of risk (*low*, *medium*, *high*, and *unacceptable*), and describe how typical design choices map to them. Our goal is not to propose a metric, but to show that one way exists to start measuring the privacy risk, and it could be based on regulatory frameworks already adopted globally.

**Low Risk.** Inference is performed locally on the device. The input is a non-identifiable data type (e.g., body keypoints, silhouettes, or RF signals). Data are not stored or are stored in encrypted form, only for short periods. Transmission is absent or encrypted. These setups follow the principles of data minimization and privacy by design (GDPR Art. 5, 25), and reflect OECD guidelines such as *Collection Limitation* and *Security Safeguards*. These systems typically fall outside the scope of regulated biometric data—such as BIPA or HIPAA, unless linked with health records.

**Medium Risk.** Input data include structured but potentially re-identifiable information (e.g., depth maps or thermal images). Inference is local, but training or aggregation may occur on external servers. Short-term storage or encrypted transfer is present. These setups may comply with regulation, but require documentation and justification under GDPR (e.g., Art. 6 on a lawful basis). As depth and thermal data may enable biometric profiling, it is crucial to comply with regulatory regimes such as the BIPA with consent and retention rules (e.g., 740 ILCS 14/15(b), (d)), HIPAA's PHI safeguards under the *Privacy* and *Security Rules*, and OECD's principles—including use limitation and accountability.

**High Risk.** RGB video is collected and stored. Inference may be offloaded to cloud services. Data are transmitted without strong encryption or stored long-term. The system is used in sensitive contexts such as health or workplace monitoring. This configuration may involve biometric data (GDPR Art. 9) and fall under high-risk use cases in the AI Act (Annex III), requiring a full risk management process. In the U.S., health-linked biometric data become PHI under HIPAA's Privacy/Security Rules (including de-identification rules under §164.514(b), (c) and Safe Harbor/Expert Determination methods), and BIPA's strict notice, consent, and retention obligations (740 ILCS 14/15) also apply.

**Unacceptable Risk.** Biometric data are collected without consent. RGB or 3D mesh data are stored and reused for purposes unrelated to the original intent. Data are processed in jurisdictions without regulatory protection or oversight. These configurations likely violate GDPR and PIPL, BIPA's consent and disclosure requirements (e.g., 740 ILCS 14/15), HIPAA privacy safeguards, OECD/UN principles, and would be non-compliant even under basic legal scrutiny.

Figure 4 summarizes examples of practices in HPE, grouped by privacy risk level and aligned with relevant legal provisions. This classification does not require any new regulation. It builds directly on existing law and links technical decisions in HPE to levels of risk. These levels could be part of benchmark reporting and accuracy metrics to support fair comparisons between systems.

### 3.3 Toward practical privacy indicators

The risk levels discussed in the previous section can help assess the privacy exposure of HPE systems, but are insufficient for comparing or evaluating models in practice. Unlike accuracy, a privacy score is never reported in benchmarks or model documentation. This makes it difficult to assess the trade-off between performance and privacy or to understand how a system handles personal data.

We argue that it is possible to introduce basic privacy indicators—simple, interpretable, and aligned with legal principles—that can be used during system design and evaluation. These indicators do not aim to capture all aspects of privacy, but they can make privacy visible and comparable. Some possible approaches, which could also be combined with each other, may be based on the ideas as follows. These examples show that privacy in HPE can be evaluated in multiple ways, using system-level properties and practical indicators. None is definitive, but together they could support benchmarks where privacy is not ignored, but reported and considered alongside performance.

#### 3.3.1 Scoring input modalities

The input data used by an HPE system directly affect the amount of SPI exposed. Some modalities, such as keypoints or RF signals, contain little or no identifiable information. Others, like high-resolution RGB with visible faces and background, capture a wide range of biometric and contextual features. A scoring scheme could assign a privacy exposure value to each modality. For example:

- 0 points: keypoints, RF signals, IMUs;
- 1 point: silhouettes, depth maps, thermal images;
- 2 points: obfuscated or low-res RGB;
- 3 points: high-res RGB with visible face and background.

This type of score reflects the level of exposure before any processing, and could be extended to systems using multiple modalities.

### 3.3.2 Penalizing risky handling of data

Privacy is not determined by the input alone. It also depends on what happens to the data after they are captured. Design choices such as where inference occurs, how data are transmitted, and whether they are stored introduce additional levels of risk. One way to consider this aspect could be to design a penalty system that assigns one point to each design choice that increases privacy risk:

- cloud-based inference;
- unencrypted data transfer;
- long-term storage;
- reuse of data for different purposes.

This cumulative score would reflect how many risk factors characterize the pipeline.

### 3.3.3 Privacy labels based on legal categories

Instead of numeric scores, an HPE system could carry a structured label that considers privacy-relevant properties, inspired by GDPR and the AI Act (but also other frameworks). For example:

- uses biometric data (GDPR Art. 9): yes/no;
- AI Act risk level: minimal / limited / high;
- data retention: none / short / long;
- international transfer: none / with safeguards / unrestricted.

This label does not measure privacy but highlights legal exposure and promotes early consideration of regulatory constraints.

### 3.3.4 Re-identification resistance

A more empirical option could be based on measuring the extent to which a system's output can help re-identify individuals. This option could use, for example, a re-identification classifier to test whether blurred images, silhouettes, or depth maps still reveal identity.

Recent studies have shown that 2D/3D skeletons and 3D meshes can support human re-identification. In particular, skeleton-guided feature learning enables recognition across significant visual variations [113, 114], whereas mesh-based modeling highlights the potential of structured geometric outputs to preserve discriminative cues for re-identification in complex scenarios [115]. Various anonymization strategies have been proposed to mitigate this risk. For example, motion retargeting and adversary-guided retargeting which, respectively, anonymize skeleton data while preserving realistic motion patterns [116], and reduce the re-identification risk without degrading action utility [117]. Also, 3D human pose and shape reconstruction have demonstrated that structured outputs facilitate re-identification, highlighting the need for privacy-preserving mechanisms [118].

Beyond pose data, visual privacy scores derived from image attributes provide an empirical framework for quantifying privacy risks [119]. In addition, self-supervised learning can suppress private information at the feature level [120].

Re-identification resistance could thus help generate a privacy score based on actual model behavior rather than assumptions. It could be used alongside accuracy, latency, and other performance metrics to compare systems more fairly in privacy-sensitive settings.

## 3.4 Performance vs Privacy

Many HPE methods that improve privacy tend to reduce accuracy as a side effect. For example, using depth maps, thermal images, or radio frequency data instead of RGB often decreases keypoint detection quality. Local inference limits the model size and capacity. Obfuscation techniques, such as blurring or pixelation, can hide identifiable traits but degrade spatial precision. These trade-offs are well known, but rarely measured or reported explicitly.

Existing benchmarks typically evaluate models on clean RGB input in ideal conditions. As a result, models optimized for accuracy may perform poorly when introducing privacy constraints.

Systems designed for real-world deployment—such as rehabilitation, workplace safety, or smart environments—must be tested under realistic assumptions about sensing, storage, and processing.

We suggest that benchmarks should include both accuracy and privacy risk, based on the levels described in Section 3. For example, a model could be reported as achieving 85% AP@0.5 under a medium privacy risk configuration, or 78% under a low-risk configuration. This would make trade-offs clear and comparable.

Various directions may help reduce the gap between privacy and performance. These include designing models that are robust to low-information inputs (e.g., edge-based features, coarse silhouettes), fine-tuning on privacy-filtered datasets, and using modality fusion (e.g., combining IMU and depth). Also, privacy-aware training objectives can be explored. Lightweight architectures and real-time inference pipelines can further support local and secure deployment.

Privacy and accuracy are not mutually exclusive, but optimizing for both requires shifting how we design, train, and evaluate HPE systems.

## 4 Alternative views

Despite growing attention to privacy, much of the current HPE research is shaped by assumptions that sideline or oversimplify it. These assumptions often stem from choices in datasets, benchmarks, or evaluation protocols, where privacy is either abstracted away or treated as an afterthought. Below, we highlight alternative perspectives that continue to inform the field and which we believe should be critically examined.

**Accuracy is the main objective.** Most HPE models are designed and evaluated based on accuracy scores such as PCK, MPJPE, or AP [121–123]. This view only considers privacy when it comes to deploying models, not during development or benchmarking. This underestimates the role of upstream data flows and model behavior.

This approach overlooks real-world scenarios where privacy is a legal and ethical requirement from the very beginning. A highly accurate model that cannot be deployed due to privacy risks has a limited practical value.

**Switching to non-RGB data guarantees privacy.** Some researchers assume that using depth maps, LiDAR scans, thermal images, or RF signals is enough to guarantee privacy. HPE systems that rely on these inputs are described as "privacy-preserving" [124–126].

However, this is overstated, irrespective of the data modality. In general, non-RGB data can reveal body shape, walking style, or even signs of stress or illness. What matters is what the model can still learn or infer, rather than what the image looks like.

**Privacy is difficult to measure.** Another common belief is that privacy is too vague or context-specific to define, and that is why it is ignored [127–130]. However, some regulations define several aspects of privacy, although they do not give a universal definition of privacy.

In the previous section, we proposed dimensions along which we can assess privacy risks in distributed HPE systems. These dimensions are grounded in international regulations and include the type of data used, how the model is trained, where inference takes place, and how data are stored.

We thus believe that privacy is not impossible to measure, but requires clear and shared criteria to comply with international regulations.

**Synthetic or augmented data preserve privacy.** Some works propose generating synthetic datasets [131–133] or augmenting real datasets with generative models as a privacy workaround. This approach reduces reliance on real-world data, but introduces other risks. For example, synthetic data can still replicate sensitive patterns from training input (e.g., via memorization in GANs), and augmented data may retain identifiable traces if based on insufficiently anonymized inputs. Risk should be evaluated based on what can be reconstructed or inferred from model output.

**Privacy-preserving means not storing or transferring data.** Some system architectures are often claimed to be privacy-preserving only because they do not retain input data after inference [134–136]. Although data retention is crucial to assess privacy, it is not the only aspect. Real-time systems that process data in memory without logging may still expose personal information through live

streams, insecure APIs, or inference-time side channels. Privacy requires more than avoiding storage, it requires minimizing collection, restricting transmission, and securing processing across all layers.

These views reflect a broader trend in HPE research: privacy is often considered a deployment detail, not a fundamental design concern. We believe this is incompatible with the regulatory, ethical, and societal expectations surrounding real-world AI systems. Reframing privacy as a measurable, multi-dimensional property of system behavior (not just data type or retention policy) is crucial if HPE models are to be trusted, adopted, and legally compliant.

## 5 Conclusions

Human pose estimation (HPE) has reached high levels of accuracy and is ready for deployment in many real-world settings. But if these systems are to be used in workplaces, healthcare, rehabilitation, or public spaces, they must go beyond performance. Accuracy alone is no longer enough.

In this position paper, we have argued that privacy must be treated as a core evaluation criterion in HPE, not an afterthought. We have highlighted how current practices overlook key risk factors—such as data type, storage, and inference location—and how common assumptions, such as *non-RGB equals privacy* fail. And we have outlined how legal frameworks already provide structure for thinking about risk, while also showing that measurable indicators can be introduced into benchmarks to reflect privacy exposure.

Then, we have presented some practical directions for evaluating privacy, including modality scores, penalty-based indicators, legal labeling, and re-identification resistance testing. These tools do not replace regulation, but make privacy visible so that researchers can compare systems by how well they perform and to what extent they respect privacy.

The time has come for a shift. Privacy and trust must be treated as design requirements, and we already have the means to reduce identifiability, secure data, and design models that are compatible with regulation and societal expectations. The next step is to build the benchmarks, tools, and incentives that make privacy a standard part of the progress in HPE.

## References


[1] F. Zhang, X. Zhu, and C. Wang, "A Comprehensive Survey on Single-Person Pose Estimation in Social Robotics," *International Journal of Social Robotics*, vol. 14, no. 9, pp. 1995–2008, 2022.

[2] M. Cormier, A. Clepe, A. Specker, and J. Beyerer, "Where are we with human pose estimation in real-world surveillance?," in *2022 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)*, pp. 591–601, 2022.

[3] L. Wang, D. Q. Huynh, and P. Koniusz, "A Comparative Review of Recent Kinect-Based Action Recognition Algorithms," *IEEE Transactions on Image Processing*, vol. 29, pp. 15–28, 2020.

[4] A. Badiola-Bengoa and A. Mendez-Zorrilla, "A Systematic Review of the Application of Camera-Based Human Pose Estimation in the Field of Sport and Physical Exercise," *Sensors*, vol. 21, no. 18, 2021.

[5] J. Stenum, K. M. Cherry-Allen, C. O. Pyles, R. D. Reetzke, M. F. Vignos, and R. T. Roemmich, "Applications of Pose Estimation in Human Health and Performance across the Lifespan," *Sensors*, vol. 21, no. 21, 2021.

[6] C. Zheng, W. Wu, C. Chen, T. Yang, S. Zhu, J. Shen, N. Kehtarnavaz, and M. Shah, "Deep Learning-based Human Pose Estimation: A Survey," *ACM Computing Surveys*, vol. 56, no. 1, pp. 1–37, 2023.

[7] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, "Human3.6m: Large scale datasets and predictive methods for 3d human sensing in natural environments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 7, pp. 1325–1339, 2014.

[8] J. Liu, A. Shahroudy, M. Perez, G. Wang, L.-Y. Duan, and A. C. Kot, "NTU RGB+D 120: A Large-Scale Benchmark for 3D Human Activity Understanding," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 10, pp. 2684–2701, 2020.

[9] H. Joo, H. Liu, L. Tan, L. Gui, B. Nabbe, I. Matthews, T. Kanade, S. Nobuhara, and Y. Sheikh, "Panoptic studio: A massively multiview system for social motion capture," in *International Conference on Computer Vision (ICCV)*, 2015.

[10] M. Andriluka, L. Pishchulin, P. Gehler, and B. Schiele, "2d human pose estimation: New benchmark and state of the art analysis," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3686–3693, 2014.

[11] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common Objects in Context," in *2014 European Conference on Computer Vision (ECCV)*, pp. 740–755, Springer International Publishing, 2014.

[12] G. Varol, J. Romero, X. Martin, N. Mahmood, M. J. Black, I. Laptev, and C. Schmid, "Learning from synthetic humans," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4627–4635, 2017.

[13] K. Rasheed, A. Qayyum, M. Ghaly, A. Al-Fuqaha, A. Razi, and J. Qadir, "Explainable, trustworthy, and ethical machine learning for healthcare: A survey," *Computers in Biology and Medicine*, vol. 149, p. 106043, Oct. 2022.

[14] Y. Ben-Shabat, X. Yu, F. Saleh, D. Campbell, C. Rodriguez-Opazo, H. Li, and S. Gould, "The IKEA ASM Dataset: Understanding People Assembling Furniture through Actions, Objects and Pose," in *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 846–858, 2021.

[15] R. Aguilar-Ortega, R. Berral-Soler, I. Jiménez-Velasco, F. J. Romero-Ramírez, M. García-Marín, J. Zafra-Palma, R. Muñoz-Salinas, R. Medina-Carnicer, and M. J. Marín-Jiménez, "Uco physical rehabilitation: New dataset and study of human pose estimation methods on physical rehabilitation exercises," *Sensors*, vol. 23, p. 8862, Oct. 2023.

[16] S. An, Y. Li, and U. Ogras, "mri: multi-modal 3d human pose estimation dataset using mmwave, rgb-d, and inertial sensors," in *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22, (Red Hook, NY, USA), Curran Associates Inc., 2022.

[17] J. Chai, H. Zeng, A. Li, and E. W. Ngai, "Deep learning in computer vision: A critical review of emerging techniques and application scenarios," *Machine Learning with Applications*, vol. 6, p. 100134, Dec. 2021.

[18] S. Ravi, P. Climent-Pérez, and F. Florez-Revuelta, "A review on visual privacy preservation techniques for active and assisted living," *Multimedia Tools and Applications*, vol. 83, p. 14715–14755, July 2023.

[19] G. Zhang, B. Liu, T. Zhu, A. Zhou, and W. Zhou, "Visual privacy attacks and defenses in deep learning: a survey," *Artificial Intelligence Review*, vol. 55, p. 4347–4401, Jan. 2022.

[20] Y. Chen, Y. Tian, and M. He, "Monocular human pose estimation: A survey of deep learning-based methods," *Computer Vision and Image Understanding*, vol. 192, p. 102897, Mar. 2020.

[21] T. L. Munea, Y. Z. Jembre, H. T. Weldegebriel, L. Chen, C. Huang, and C. Yang, "The Progress of Human Pose Estimation: A Survey and Taxonomy of Models Applied in 2D Human Pose Estimation," *IEEE Access*, vol. 8, p. 133330–133348, 2020.

[22] Y. Yang, H. Zhang, A. B. Fernández, S. Alemany, S. Chen, and G. Zhang, "Digitalization of 3-D Human Bodies: A Survey," *IEEE Transactions on Consumer Electronics*, vol. 70, p. 3152–3166, Feb. 2024.

[23] Y. Tian, H. Zhang, Y. Liu, and L. Wang, "Recovering 3D Human Mesh From Monocular Images: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, p. 15406–15425, Dec. 2023.

[24] C. Zheng, W. Wu, C. Chen, T. Yang, S. Zhu, J. Shen, N. Kehtarnavaz, and M. Shah, "Deep Learning-based Human Pose Estimation: A Survey," *ACM Computing Surveys*, vol. 56, p. 1–37, Aug. 2023.

[25] W. Liu, Q. Bao, Y. Sun, and T. Mei, "Recent Advances of Monocular 2D and 3D Human Pose Estimation: A Deep Learning Perspective," *ACM Computing Surveys*, vol. 55, p. 1–41, Nov. 2022.

[26] G. Lan, Y. Wu, F. Hu, and Q. Hao, "Vision-Based Human Pose Estimation via Deep Learning: A Survey," *IEEE Transactions on Human-Machine Systems*, vol. 53, p. 253–268, Feb. 2023.

[27] R. B. Neupane, K. Li, and T. F. Boka, "A survey on deep 3D human pose estimation," *Artificial Intelligence Review*, vol. 58, Nov. 2024.

[28] B. Rajkumar and S. Satish Narayana, *Fog and Edge Computing: Principles and Paradigms*. Wiley, Jan. 2019.

[29] A. Shahroudy, J. Liu, T.-T. Ng, and G. Wang, "Ntu rgb+d: A large scale dataset for 3d human activity analysis," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, June 2016.

[30] J. Liu, A. Shahroudy, M. Perez, G. Wang, L.-Y. Duan, and A. C. Kot, "Ntu rgb+d 120: A large-scale benchmark for 3d human activity understanding," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, p. 2684–2701, Oct. 2020.

[31] J. Li, J. Zhang, Z. Wang, S. Shen, C. Wen, Y. Ma, L. Xu, J. Yu, and C. Wang, "LiDARCap: Long-range Marker-less 3D Human Motion Capture with LiDAR Point Clouds," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, pp. 20502–20512, 2022.

[32] G. Varol, J. Romero, X. Martin, N. Mahmood, M. J. Black, I. Laptev, and C. Schmid, "Learning from Synthetic Humans," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2017.

[33] J. Yang, H. Huang, Y. Zhou, X. Chen, Y. Xu, S. Yuan, H. Zou, C. X. Lu, and L. Xie, "MM-Fi: Multi-Modal Non-Intrusive 4D Human Dataset for Versatile Wireless Sensing," in *Thirty-seventh Conf. on Neural Inf. Process. Systems Datasets and Benchmarks Track*, 2023.

[34] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, "Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 7, p. 1325–1339, 2014.

[35] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand, "Capturing the human figure through a wall," *ACM Trans. on Graph.*, vol. 34, no. 6, p. 1–13, 2015.

[36] M. Zhao, T. Li, M. A. Alsheikh, Y. Tian, H. Zhao, A. Torralba, and D. Katabi, "Through-Wall Human Pose Estimation Using Radio Signals," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, 2018.

[37] M. Zhao, Y. Tian, H. Zhao, M. A. Alsheikh, T. Li, R. Hristov, Z. Kabelac, D. Katabi, and A. Torralba, "RF-based 3D skeletons," in *Conf. of the ACM Special Interest Group on Data Commun.*, 2018.

[38] K. Meng and Y. Meng, "Through-Wall Pose Imaging in Real-Time with a Many-to-Many Encoder/Decoder Paradigm," in *IEEE Int. Conf. On Mach. Learn. And Appl.*, 2019.

[39] A. Sengupta, F. Jin, and S. Cao, "NLP based Skeletal Pose Estimation using mmWave Radar Point-Cloud: A Simulation Approach," in *2020 IEEE Radar Conf.*, 2020.

[40] A. Sengupta, F. Jin, R. Zhang, and S. Cao, "mm-Pose: Real-Time Human Skeletal Posture Estimation Using mmWave Radars and CNNs," *IEEE Sensors J.*, vol. 20, no. 17, p. 10032–10044, 2020.

[41] A. Sengupta and S. Cao, "mmPose-NLP: A Natural Language Processing Approach to Precise Skeletal Pose Estimation Using mmWave Radars," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 11, p. 8418–8429, 2023.

[42] C. Yang, X. Wang, and S. Mao, "RFID-Pose: Vision-Aided Three-Dimensional Human Pose Estimation With Radio-Frequency Identification," *IEEE Trans. Rel.*, vol. 70, no. 3, p. 1218–1231, 2021.

[43] C. Yang, X. Wang, and S. Mao, "Subject-adaptive Skeleton Tracking with RFID," in *Int. Conf. on Mobility, Sens. and Networking*, 2020.

[44] C. Yang, L. Wang, X. Wang, and S. Mao, "Environment Adaptive RFID-Based 3D Human Pose Tracking With a Meta-Learning Approach," *IEEE J. of Radio Freq. Identification*, vol. 6, p. 413–425, 2022.

[45] C. Yang, Z. Wang, and S. Mao, "RFPose-GAN: Data Augmentation for RFID based 3D Human Pose Tracking," in *Int. Conf. on RFID Technol. and Appl.*, 2022.

[46] C. Yu, D. Zhang, Z. Wu, C. Xie, Z. Lu, Y. Hu, and Y. Chen, "MobiRFPose: Portable RF-Based 3D Human Pose Camera," *IEEE Trans. Multimedia*, vol. 26, p. 3715–3727, 2024.

[47] C. Xie, D. Zhang, Z. Wu, C. Yu, Y. Hu, and Y. Chen, "RPM: RF-Based Pose Machines," *IEEE Trans. Multimedia*, vol. 26, p. 637–649, 2024.

[48] Y. Cao, A. Dhekne, and M. Ammar, "ViSig: Automatic Interpretation of Visual Body Signals Using On-Body Sensors," *ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 7, no. 1, p. 1–27, 2023.

[49] F. Wang, S. Panev, Z. Dai, J. Han, and D. Huang, "Can WiFi Estimate Person Pose?," 2019.

[50] F. Wang, S. Zhou, S. Panev, J. Han, and D. Huang, "Person-in-WiFi: Fine-Grained Person Perception Using WiFi," in *IEEE/CVF Int. Conf. on Computer Vision*, 2019.

[51] L. Guo, Z. Lu, X. Wen, S. Zhou, and Z. Han, "From Signal to Image: Capturing Fine-Grained Human Poses With Commodity Wi-Fi," *IEEE Commun. Lett.*, vol. 24, no. 4, p. 802–806, 2020.

[52] Y. Wang, L. Guo, Z. Lu, X. Wen, S. Zhou, and W. Meng, "From Point to Space: 3D Moving Human Pose Estimation Using Commodity WiFi," *IEEE Commun. Lett.*, vol. 25, no. 7, p. 2235–2239, 2021.

[53] J. Yang, Y. Zhou, H. Huang, H. Zou, and L. Xie, "MetaFi: Device-Free Pose Estimation via Commodity WiFi for Metaverse Avatar Simulation," in *World Forum on Internet of Things*, 2022.

[54] Y. Ren, Z. Wang, Y. Wang, S. Tan, Y. Chen, and J. Yang, "3D Human Pose Estimation Using WiFi Signals," in *ACM Conf. on Embedded Networked Sensor Syst.*, 2021.

[55] J. Geng, D. Huang, and F. De la Torre, "DensePose From WiFi," 2023.

[56] Y. Zhou, H. Huang, S. Yuan, H. Zou, L. Xie, and J. Yang, "MetaFi++: WiFi-Enabled Transformer-Based Human Pose Estimation for Metaverse Avatar Simulation," *IEEE Internet of Things J.*, vol. 10, no. 16, p. 14128–14136, 2023.

[57] Y.-C. Chen, Z.-K. Huang, L. Pang, J.-Y. Jiang-Lin, C.-H. Kuo, H.-H. Shuai, and W.-H. Cheng, "Seeing the unseen: Wifi-based 2D human pose estimation via an evolving attentive spatial-Frequency network," *Pattern Recognition Lett.*, vol. 171, p. 21–27, 2023.

[58] C. Tang, W. Li, S. Vishwakarma, F. Shi, S. Julier, and K. Chetty, "MDPose: Human Skeletal Motion Reconstruction Using WiFi Micro-Doppler Signatures," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 60, no. 1, p. 157–167, 2024.

[59] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-time human pose recognition in parts from single depth images," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2011.

[60] H. Yub Jung, S. Lee, Y. Seok Heo, and I. Dong Yun, "Random Tree Walk Toward Instantaneous 3D Human Pose Estimation," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2015.

[61] A. Haque, B. Peng, Z. Luo, A. Alahi, S. Yeung, and L. Fei-Fei, *Towards Viewpoint Invariant 3D Human Pose Estimation*, pp. 160–177. Springer International Publishing, 2016.

[62] K. Wang, S. Zhai, H. Cheng, X. Liang, and L. Lin, "Human Pose Estimation from Depth Images via Inference Embedded Multi-task Learning," in *ACM International Conference on Multimedia*, 2016.

[63] M. J. Marín-Jiménez, F. J. Romero-Ramirez, R. Muñoz-Salinas, and R. Medina-Carnicer, "3D human pose estimation from depth maps using a deep combination of poses," *J. of Visual Communication and Image Representation*, vol. 55, p. 627–639, 2018.

[64] V. Srivastav, A. Gangi, and N. Padoy, *Human Pose Estimation on Privacy-Preserving Low-Resolution Depth Images*, p. 583–591. Springer International Publishing, 2019.

[65] A. Martiez-Gonzalez, M. Villamizar, O. Canevet, and J.-M. Odobez, "Real-time Convolutional Networks for Depth-based Human Pose Estimation," in *IEEE/RSJ Int. Conf. on Intell. Robots and Systems*, 2018.

[66] J. Y. Chang, G. Moon, and K. M. Lee, "V2V-PoseNet: Voxel-to-Voxel Prediction Network for Accurate 3D Hand and Human Pose Estimation from a Single Depth Map," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, 2018.

[67] Y. Guo, Z. Li, Z. Li, X. Du, S. Quan, and Y. Xu, "PoP-Net: Pose Over Parts Network for Multi-Person 3D Pose Estimation From a Depth Image," in *IEEE/CVF Winter Conf. on Appl. of Computer Vision*, pp. 1240–1249, 2022.

[68] Y. Zhou, H. Dong, and A. El Saddik, "Learning to Estimate 3D Human Pose From Point Cloud," *IEEE Sensors J.*, vol. 20, no. 20, p. 12334–12342, 2020.

[69] Z. Zhang, L. Hu, X. Deng, and S. Xia, "Weakly supervised adversarial learning for 3D human pose estimation from point clouds," *IEEE Trans. Vis. Comput. Graphics*, vol. 26, no. 5, pp. 1851–1859, 2020.

[70] Z. Zhang, L. Hu, X. Deng, and S. Xia, "Sequential 3D Human Pose Estimation Using Adaptive Point Cloud Sampling Strategy.," in *Int. Joint Conf. on Artif. Intell.*, pp. 1330–1337, 2021.

[71] K. Wang, J. Xie, G. Zhang, L. Liu, and J. Yang, "Sequential 3D Human Pose and Shape Estimation From Point Clouds," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, 2020.

[72] A. D'Eusanio, S. Pini, G. Borghi, R. Vezzani, and R. Cucchiara, "RefiNet: 3D Human Pose Refinement with Depth Maps," in *Int. Conf. on Pattern Recognition*, 2021.

[73] A. D'Eusanio, A. Simoni, S. Pini, G. Borghi, R. Vezzani, and R. Cucchiara, "Depth-based 3D human pose refinement: Evaluating the refinet framework," *Pattern Recognition Lett.*, vol. 171, p. 185–191, 2023.

[74] G. Borghi, M. Fabbri, R. Vezzani, S. Calderara, and R. Cucchiara, "Face-from-Depth for Head Pose Estimation on Depth Images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 3, p. 596–609, 2020.

[75] F. Xiong, B. Zhang, Y. Xiao, Z. Cao, T. Yu, J. T. Zhou, and J. Yuan, "A2J: Anchor-to-Joint Regression Network for 3D Articulated Pose Estimation From a Single Depth Image," in *IEEE/CVF Int. Conf. on Computer Vision*, 2019.

[76] N. Garau, N. Bisagno, P. Bródka, and N. Conci, "DECA: Deep Viewpoint-Equivariant Human Pose Estimation Using Capsule Autoencoders," in *IEEE/CVF Int. Conf. on Computer Vision*, pp. 11677–11686, 2021.

[77] Y. Ren, C. Zhao, Y. He, P. Cong, H. Liang, J. Yu, L. Xu, and Y. Ma, "LiDAR-aid Inertial Poser: Large-scale Human Motion Capture by Sparse Inertial and LiDAR Sensors," *IEEE Trans. Vis. Comput. Graphics*, vol. 29, no. 5, p. 2337–2347, 2023.

[78] Y. Ren, X. Han, C. Zhao, J. Wang, L. Xu, J. Yu, and Y. Ma, "LiveHPS: LiDAR-based Scene-level Human Pose and Shape Estimation in Free Environment," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, pp. 1281–1291, 2024.

[79] Z. Cai, L. Pan, C. Wei, W. Yin, F. Hong, M. Zhang, C. C. Loy, L. Yang, and Z. Liu, "PointHPS: Cascaded 3D Human Pose and Shape Estimation from Point Clouds," 2023.

[80] X. Wu, H. Zhang, C. Kong, Y. Wang, Y. Ju, and C. Zhao, "LiDAR-Based 3-D Human Pose Estimation and Action Recognition for Medical Scenes," *IEEE Sensors J.*, vol. 24, no. 9, p. 15531–15539, 2024.

[81] D. Ye, Y. Xie, W. Chen, Z. Zhou, L. Ge, and H. Foroosh, "LPFormer: LiDAR Pose Estimation Transformer with Multi-Task Network," 2023.

[82] D. Ye, Z. Zhou, W. Chen, Y. Xie, Y. Wang, P. Wang, and H. Foroosh, "LidarMultiNet: Towards a Unified Multi-Task Network for LiDAR Perception," 2022.

[83] A. A. Sarawade and N. N. Charniya, "Infrared Thermography and its Applications: A Review," in *Int. Conf. on Commun. and Electron. Syst.*, 2018.

[84] V. Kosonogov, L. De Zorzi, J. Honoré, E. S. Martínez-Velázquez, J.-L. Nandrino, J. M. Martinez-Selva, and H. Sequeira, "Facial thermal variations: A new marker of emotional arousal," *PLOS ONE*, vol. 12, no. 9, 2017.

[85] H. Takahashi, K. Oiwa, and A. Nozawa, "Evaluation of the effects of cold and hot environmental temperatures on the spatial distribution of facial skin temperature," in *Int. Conf. on Intell. Informat. and Biomed. Sci.*, vol. 7, pp. 374–377, 2022.

[86] K. Nishi, M. Demura, J. Miura, and S. Oishi, "Use of Thermal Point Cloud for Thermal Comfort Measurement and Human Pose Estimation in Robotic Monitoring," in *IEEE Int. Conf. on Computer Vision Workshops*, 2017.

[87] Y. Zhu, W. Lu, R. Zhang, R. Wang, and D. Robbins, "Dual-channel cascade pose estimation network trained on infrared thermal image and groundtruth annotation for real-time gait measurement," *Medical Image Anal.*, vol. 79, 2022.

[88] J. Smith, P. Loncomilla, and J. Ruiz-Del-Solar, "Human Pose Estimation Using Thermal Images," *IEEE Access*, vol. 11, p. 35352–35370, 2023.

[89] Y. Guo, Y. Chen, J. Deng, S. Li, and H. Zhou, "Identity-Preserved Human Posture Detection in Infrared Thermal Images: A Benchmark," *Sensors*, vol. 23, no. 1, 2023.

[90] G. Vdoviak and T. Sledevič, "Enhancing Keypoint Detection in Thermal Images: Optimizing Loss Function and Real-time Processing with YOLOv8n-Pose," in *Workshop on Advances in Inf. Electron. and Elect. Eng.*, 2024.

[91] I.-C. Chen, C.-J. Wang, C.-K. Wen, and S.-J. Tzou, "Multi-Person Pose Estimation Using Thermal Images," *IEEE Access*, vol. 8, p. 174964–174971, 2020.

[92] M. Lupión, V. González-Ruiz, J. Medina-Quero, J. F. Sanjuan, and P. M. Ortigosa, "THPoseLite, a Lightweight Neural Network for Detecting Pose in Thermal Images," *IEEE Internet of Things J.*, vol. 10, no. 17, p. 15060–15073, 2023.

[93] M. Gochoo, T.-H. Tan, F. Alnajjar, J.-W. Hsieh, and P.-Y. Chen, "Lownet: Privacy Preserved Ultra-Low Resolution Posture Image Classification," in *IEEE Int. Conf. on Image Process.*, 2020.

[94] M. Zhao, S. Yue, D. Katabi, T. S. Jaakkola, and M. T. Bianchi, "Learning sleep stages from radio signals: a conditional adversarial architecture," in *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, ICML'17, p. 4100–4109, JMLR.org, 2017.

[95] M. Zhao, T. Li, M. A. Alsheikh, Y. Tian, H. Zhao, A. Torralba, and D. Katabi, "Through-wall human pose estimation using radio signals," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7356–7365, 2018.

[96] M. Zhao, F. Adib, and D. Katabi, "Emotion recognition using wireless signals," *Commun. ACM*, vol. 61, p. 91–100, Aug. 2018.

[97] J. Liu and L. Zhang, "Indoor Privacy-preserving Action Recognition via Partially Coupled Convolutional Neural Network," in *2020 Int. Conf. on Artif. Intell. and Computer Eng.*, pp. 292–295, 2020.

[98] M. Ryoo, B. Rothrock, C. Fleming, and H. Yang, "Privacy-preserving human activity recognition from extreme low resolution," in *Proc. of the Thirty-First AAAI Conf. on Artif. Intell.*, pp. 4255–4262, 2017.

[99] C. Wang, F. Zhang, X. Zhu, and S. S. Ge, "Low-resolution human pose estimation," *Pattern Recognition*, vol. 126, 2022.

[100] P. Agrawal and P. J. Narayanan, "Person De-Identification in Videos," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 3, pp. 299–310, 2011.

[101] D. Chen, Y. Chang, R. Yan, and J. Yang, "Tools for Protecting the Privacy of Specific Individuals in Video," *EURASIP J. on Advances in Signal Process.*, vol. 2007, no. 1, 2007.

[102] J. R. Padilla-López, A. A. Chaaraoui, and F. Flórez-Revuelta, "Visual privacy protection methods," *Expert Syst. Appl.*, vol. 42, no. 9, p. 4177–4195, 2015.

[103] S. Ahmad, P. Morerio, and A. Del Bue, "Event Anonymization: Privacy-Preserving Person Re-Identification and Pose Estimation in Event-Based Vision," *IEEE Access*, vol. 12, pp. 66964–66980, 2024.

[104] T. Cao, M. A. Armin, S. Denman, L. Petersson, and D. Ahmedt-Aristizabal, "In-Bed Human Pose Estimation from Unseen and Privacy-Preserving Image Domains," in *2022 IEEE 19th Int. Symp. on Biomed. Imag.*, 2022.

[105] V. Srivastav, A. Gangi, and N. Padoy, "Human Pose Estimation on Privacy-Preserving Low-Resolution Depth Images," in *Medical Image Comput. and Computer Assisted Intervention*, p. 583–591, Springer-Verlag, 2019.

[106] C. Hinojosa, J. C. Niebles, and H. Arguello, "Learning Privacy-preserving Optics for Human Pose Estimation," in *2021 IEEE/CVF Int. Conf. on Computer Vision*, pp. 2553–2562, 2021.

[107] W. Huang, Y. Ni, A. Rezvani, S. Jeong, H. Chen, Y. Liu, F. Wen, and M. Imani, "Recoverable Anonymization for Pose Estimation: A Privacy-Enhancing Approach," 2024.

[108] W. Zhuang, J. Xu, C. Chen, J. Li, and L. Lyu, "COALA: A Practical and Vision-Centric Federated Learning Platform," in *Int. Conf. on Mach. Learn.*, 2024.

[109] Z. Zhang, L. Wan, W. Xu, and S. Wang, "Low-resolution human pose estimation and action recognition via pose-driven super-resolution reconstruction," *Machine Learning*, vol. 114, Apr. 2025.

[110] A. A. A. Osman, T. Bolkart, and M. J. Black, "STAR: A sparse trained articulated human body regressor," in *European Conference on Computer Vision (ECCV)*, pp. 598–613, 2020.

[111] G. Pavlakos, V. Choutas, N. Ghorbani, T. Bolkart, A. A. A. Osman, D. Tzionas, and M. J. Black, "Expressive body capture: 3D hands, face, and body from a single image," in *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 10975–10985, 2019.

[112] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black, "SMPL: A Skinned Multi-Person Linear Model," *ACM Trans. Graph.*, vol. 34, no. 6, pp. 248:1–248:16, 2015.

[113] J. Liu, B. Ni, Y. Yan, P. Zhou, S. Cheng, and J. Hu, "Pose transferrable person re-identification," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4099–4108, 2018.

[114] C. Song, Y. Huang, W. Ouyang, and L. Wang, "Mask-guided contrastive attention model for person re-identification," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1179–1188, 2018.

[115] J. Chen, X. Jiang, F. Wang, J. Zhang, F. Zheng, X. Sun, and W.-S. Zheng, "Learning 3d shape feature for texture-insensitive person re-identification," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8142–8151, 2021.

[116] S. Hanisch, J. Todt, and T. Strufe, "Pantomime: Motion data anonymization using foundation motion models," 2025.

[117] T. Carr, D. Xu, and A. Lu, "Adversary-guided motion retargeting for skeleton anonymization," 2024.

[118] S. Moon, M. Kim, Z. Qin, Y. Liu, and D. Kim, "Anonymization for skeleton action recognition," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, p. 15028–15036, June 2023.

[119] T. Orekondy, B. Schiele, and M. Fritz, "Towards a visual privacy advisor: Understanding and predicting privacy risks in images," in *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 3706–3715, 2017.

[120] I. R. Dave, C. Chen, and M. Shah, "Spact: Self-supervised privacy preservation for action recognition," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20132–20141, 2022.

[121] Y. Desmarais, D. Mottet, P. Slangen, and P. Montesinos, "A review of 3d human pose estimation algorithms for markerless motion capture," *Computer Vision and Image Understanding*, vol. 212, p. 103275, Nov. 2021.

[122] J. Wang, S. Tan, X. Zhen, S. Xu, F. Zheng, Z. He, and L. Shao, "Deep 3d human pose estimation: A review," *Computer Vision and Image Understanding*, vol. 210, p. 103225, Sept. 2021.

[123] T. Kalampokas, S. Krinidis, V. Chatzis, and G. A. Papakostas, "Performance benchmark of deep learning human pose estimation for uavs," *Machine Vision and Applications*, vol. 34, Sept. 2023.

[124] E. Martini, M. Boldo, S. Aldegheri, N. Vale, M. Filippetti, N. Smania, M. Bertucco, A. Picelli, and N. Bombieri, "Preserving data privacy and accuracy of human pose estimation software based on cnn s for remote gait analysis," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine amp; Biology Society (EMBC)*, p. 3468–3471, IEEE, July 2022.

[125] X. Yan, Y. Xu, C. Chen, and S. Zhang, "Privacy preserving for ai-based 3d human pose recovery and retargeting," *ISA Transactions*, vol. 141, p. 132–142, Oct. 2023.

[126] X. Li, Y. Zhang, I. Marsic, and R. S. Burd, *Privacy Preserving Dynamic Room Layout Mapping*, p. 61–70. Springer International Publishing, 2016.

[127] B. P. Knijnenburg, X. Page, P. Wisniewski, H. R. Lipford, N. Proferes, and J. Romano, *Introduction and Overview*, p. 1–11. Springer International Publishing, 2022.

[128] B. C. Stahl, D. Schroeder, and R. Rodrigues, *Privacy*, p. 25–37. Springer International Publishing, Nov. 2022.

[129] A. Ziller, T. T. Mueller, R. Braren, D. Rueckert, and G. Kaissis, "Privacy: An axiomatic approach," *Entropy*, vol. 24, p. 714, May 2022.

[130] A. Seppälä, P. Nykänen, and P. Ruotsalainen, "Privacy-related context information for ubiquitous health," *JMIR mhealth and uhealth*, vol. 2, p. e12, Mar. 2014.

[131] Z. Qian, T. Callender, B. Cebere, S. M. Janes, N. Navani, and M. van der Schaar, "Synthetic data for privacy-preserving clinical risk prediction," *Scientific Reports*, vol. 14, Oct. 2024.

[132] T. Dayarathna, T. Muthukumarana, Y. Rathnayaka, S. Denman, C. de Silva, A. Pemasiri, and D. Ahmedt-Aristizabal, "Privacy-preserving in-bed pose monitoring: A fusion and reconstruction study," *Expert Systems with Applications*, vol. 213, p. 119139, Mar. 2023.

[133] S. Juraev, A. Ghimire, J. Alikhanov, V. Kakani, and H. Kim, "Exploring human pose estimation and the usage of synthetic data for elderly fall detection in real-world surveillance," *IEEE Access*, vol. 10, p. 94249–94261, 2022.

[134] M. Fernández, J. Jaimunk, and B. Thuraisingham, "A privacy-preserving architecture and data-sharing model for cloud-iot applications," *IEEE Transactions on Dependable and Secure Computing*, vol. 20, p. 3495–3507, July 2023.

[135] A. Gkoulalas-Divanis, G. Loukides, and J. Sun, "Publishing data from electronic health records while preserving privacy: A survey of algorithms," *Journal of Biomedical Informatics*, vol. 50, p. 4–19, Aug. 2014.

[136] L. Dutkiewicz, Y. Miadzvetskaya, H. Ofe, A. Barnett, L. Helminger, S. Lindstaedt, and A. Trügler, *Privacy-Preserving Techniques for Trustworthy Data Sharing: Opportunities and Challenges for Future Research*, p. 319–335. Springer International Publishing, 2022.