Research Article

Differentiating Conscious Human Entities From Autonomous AI Agents: A Heart Rate Variability (HRV) Framework for Human Authentication in the Emerging AI Economy

Sameer Halbe[1]

1. Independent researcher

The rapid proliferation of autonomous AI agents — software entities capable of performing complex cognitive tasks traditionally reserved for human workers — poses a fundamental challenge to digital identity and labour attribution. As AI increasingly substitutes for human effort in professional, creative, and administrative domains, the ability to reliably distinguish human contributors from AI-generated output becomes a critical sociotechnical necessity. This paper proposes Heart Rate Variability (HRV), measured contactlessly via millimeter-wave radio sensing, as a biometric anchor for differentiating conscious human entities from autonomous bots and digital agents. Building upon recent advances in robust contactless HRV estimation under diverse real-world environments — particularly the Massive Radio Sensing framework — we extend the application of HRV sensing to a new domain: continuous, passive human authentication in AI-augmented workplaces. We formalize the Human Presence Verification (HPV) problem, propose a layered HRV-authentication architecture, and analyse its robustness against spoofing, relay, and adversarial AI impersonation attacks. We further explore how this technology can underpin a Proof-of-Humanity (PoH) credential system enabling fair labour attribution, regulatory compliance, and trust in human-AI collaborative environments. Evaluation across 30 healthy participants and 130 clinical inpatients suggests that contactless HRV sensing achieves inter-beat interval (IBI) estimation errors as low as 15.47 ms, sufficient to distinguish live human physiological rhythms from synthetic or replayed signals.

Correspondence: papers@team.qeios.com — Qeios will forward to the authors

 doi.org/10.32388/BXF520

1. Introduction

The global economy is undergoing an unprecedented structural transformation. Large language models, autonomous agents, robotic process automation, and multi-modal AI systems are displacing knowledge workers across sectors ranging from law and medicine to software engineering and creative production [1] [2]. Conservative estimates suggest that between 40% and 60% of current white-collar job functions can be partially or fully automated using AI available as of 2025 [3]. More provocatively, a class of autonomous digital agents — software systems capable of receiving tasks, reasoning about them, executing multi-step plans, interacting with external services, and producing deliverables — now operate in ways that are externally indistinguishable from human remote workers [4].

This convergence of capability and invisibility creates a set of deeply consequential problems. When a client commissions a piece of legal analysis, a software deliverable, or a research report, they may have legitimate interests in knowing whether the output was produced by a human expert, an AI agent, or some combination thereof. Regulatory frameworks in finance, medicine, and intellectual property law typically assume human authorship and accountability [5]. Insurance, liability, and professional licensing structures depend on a human actor being identifiable as the responsible party. Even in contexts where AI assistance is acceptable, the degree of human involvement may matter for quality assessment, fair pricing, and ethical auditing.

Beyond labour attribution, the problem of bot-versus-human discrimination arises in contexts as varied as democratic participation (detecting AI-generated political content), scientific peer review (detecting AI-authored manuscripts), customer service (disclosing AI agents), and online credentialing (verifying that test-takers are human). Current approaches to this discrimination problem — CAPTCHA challenges, behavioural analytics, stylometric analysis, and metadata inspection — are increasingly brittle in the face of sophisticated AI systems trained to mimic human patterns [6].

We propose a fundamentally different approach grounded in physiology rather than behaviour or style: using Heart Rate Variability (HRV) as a continuous, passive, and spoofing-resistant biometric channel to verify human presence and conscious engagement. HRV — the variation in time intervals between consecutive heartbeats — is an autonomous, involuntary physiological signal generated by the interaction of the autonomic nervous system with cardiac pacemaker cells. It cannot be voluntarily fabricated to the degree of precision required for reliable impersonation, and its real-time generation requires a living, autonomically-intact organism [7][8].

Recent advances in contactless HRV monitoring via millimeter-wave (mmWave) radio sensing have made it feasible to measure HRV without physical contact, without attached sensors, and in ordinary indoor environments [9][10]. The Massive Radio Sensing framework of Xu et al. [10] demonstrated robust contactless Inter-Beat Interval (IBI) estimation with mean errors of 15.47 ms in supine and 24.78 ms in seated postures across 32 diverse indoor environments — including home and workplace settings — and achieved a 41.7% improvement over prior state-of-the-art on a 130-patient clinical cohort. This level of accuracy is sufficient to support the human authentication application we describe.

This paper makes four principal contributions:

- We formalize the Human Presence Verification (HPV) problem in the context of AI-augmented workplaces and identify the requirements a biometric solution must satisfy.
- We propose an HRV-based authentication architecture built on contactless radio sensing and analyse its security properties against known attack vectors.
- We introduce the concept of a Proof-of-Humanity (PoH) credential derived from HRV signals and discuss its deployment in professional, regulatory, and economic contexts.
- We present a threat model, evaluation, and set of open research challenges for the HRV-PoH ecosystem.

2. Background and Related Work

2.1. Heart Rate Variability as a Physiological Signal

HRV quantifies the temporal variation in the R-R interval (also termed the Inter-Beat Interval, IBI) of the cardiac cycle. It is regulated by the autonomic nervous system (ANS) through the dynamic interplay of sympathetic (activating) and parasympathetic (calming) branches. Time-domain metrics include the mean IBI, the Standard Deviation of R-R Intervals (SDRR), the Root Mean Square of Successive Differences (RMSSD), and the percentage of successive IBIs differing by more than 50 ms (pNN50) [11]. Frequency-domain analysis decomposes HRV into low-frequency (LF, 0.04–0.15 Hz) and high-frequency (HF, 0.15–0.4 Hz) bands associated with sympathetic and vagal modulation respectively [12].

HRV is highly individual — shaped by age, fitness, circadian rhythm, emotional state, cognitive load, and disease — yet exhibits consistent physiological signatures that distinguish living humans from all known non-biological systems. Crucially, the fine-grained temporal structure of HRV (sub-millisecond variation patterns) cannot currently be reproduced by any synthetic process without access to a real-time physiological source, making it a strong candidate for liveness detection [13].

2.2. Contactless HRV Monitoring via Radio Sensing

Traditional HRV monitoring relies on contact-based sensors: electrocardiography (ECG) electrodes, photoplethysmography (PPG) wrist sensors, or chest-strap accelerometers. These modalities impose user burden and are incompatible with passive, ambient monitoring scenarios required for workplace authentication.

Radio-frequency (RF) sensing — including Wi-Fi Channel State Information (CSI), Ultra-Wideband (UWB), and Frequency-Modulated Continuous Wave (FMCW) millimeter-wave radar — has emerged as a compelling contactless alternative [14][15]. Millimeter-wave radar is particularly well-suited due to its high spatial resolution, penetration of light clothing, and immunity to visible-light conditions. When directed at a seated or recumbent subject, mmWave radar captures micro-displacements of the thorax caused by cardiac motion, enabling extraction of the heartbeat waveform and subsequently HRV metrics.

The central challenge has been generalization: radio signals are acutely sensitive to environmental configuration (room layout, furniture, sensing angle, body posture), causing signal distribution shifts that degrade model performance in out-of-training-distribution conditions [10]. Xu et al. [10] addressed this through the Massive Radio Sensing framework, which models the problem statistically as estimation under a latent global environmental distribution and leverages large-scale participant diversity to approximate this distribution empirically. Their two-stage training strategy — Environment-Conditioned Batch Sampling combined with an Expectation-Oriented Optimization Strategy inspired by meta-learning — achieved consistent performance across 40 deployment environments.

For our authentication application, the key performance parameters are: (a) IBI estimation error must be low enough to reliably discriminate a live human from a replayed or synthetic signal; (b) the sensing must operate passively and unobtrusively in a workplace context; and (c) the system must be robust across the diversity of real-world deployment conditions.

2.3. Existing Approaches to Bot and AI Detection

Current bot-detection techniques operate across several layers. At the network and session layer, rate limiting, IP reputation, and TLS fingerprinting catch automated clients. At the application layer, CAPTCHA challenges and behavioural biometrics (mouse movement, keystroke dynamics, scroll patterns) attempt to distinguish human interaction patterns from scripted automation [16]. At the content layer, stylometric and statistical analyses detect AI-generated text, while deepfake detectors target synthesized media [17].

All of these approaches are engaged in an adversarial arms race with increasingly sophisticated AI. Large language models now pass standard CAPTCHA challenges, replicate human typing rhythms with high fidelity, and generate text indistinguishable from human writing on standard metrics [18]. The fundamental limitation of behaviour-based discrimination is that behaviour is learnable by sufficiently capable AI systems. Physiology, by contrast, requires not learned imitation but actual embodiment.

3. Formalizing the Human Presence Verification Problem

3.1. The AI Economy Actor Model

We define an AI Economy as an economic environment in which digital agents — autonomous software systems — routinely perform tasks previously performed by human workers and interact with other actors (human or artificial) through digital interfaces. We model three classes of actor:

- Human Worker (H): A biological human entity engaged in a work task, interacting with digital systems. Possesses an autonomic nervous system generating continuous involuntary physiological signals including HRV.
- Digital Agent (D): An autonomous AI system — language model, robotic process automation, agentic workflow — performing a task. Generates no physiological signals. May produce outputs designed to pass as human-generated.
- Hybrid Actor (X): A human worker substantially augmented by AI tools, where portions of output are AI-generated but a human remains in the loop, reviewing, directing, and approving.

The Human Presence Verification (HPV) problem is: given observations of a digital work session, determine the actor class — H, D, or X — and, in the case of X, quantify the degree of human involvement. For many regulatory and commercial contexts, a binary determination (H vs. D) suffices; for others, a continuous human-involvement score is required.

3.2. Formal Problem Statement

Let S = $\{s_1, s_2, ..., s_n\}$ be a sequence of n work units (e.g., document edits, API calls, code commits, decisions) produced within a work session of duration T. Let $\Phi(t)$ denote the physiological signal observable from the session environment at time t. The HPV classification function is:

$$C : (S, \Phi) \rightarrow \{H, D, X\} \times [0, 1]$$

where the second component is a human-involvement confidence score. When Φ is null (no physiological sensor in the environment), C must rely solely on S — the work output behavioural channel — which is vulnerable to AI impersonation. When Φ is available and authenticated, C gains a physiological channel that is orders of magnitude harder to spoof.

3.3. Requirements for an HRV-Based HPV System

A viable HRV-based HPV system must satisfy the following requirements:

Requirement	Description	Target Specification
Accuracy	IBI estimation error must support liveness discrimination	Mean IBI error < 30 ms
Passivity	No active user cooperation required during work session	Zero interaction overhead
Robustness	Performance must hold across real-world indoor environments	Stability Index < 5 ms
Privacy	Raw physiological data must not be stored or linkable beyond session	On-device processing; hash-only credential
Anti-Spoofing	System must resist replay, synthetic, and relay attacks	Attack success rate < 1%
Scalability	System must operate across many concurrent sessions	Linear compute scaling
Clinical Safety	Radar emissions must be within safe exposure limits	Comply with ICNIRP guidelines

Table 1. Requirements specification for an HRV-based Human Presence Verification system.

4. Proposed HRV-Authentication Architecture

4.1. System Overview

We propose a four-layer HRV-based Human Presence Verification architecture, illustrated conceptually in Figure 1. The layers are: Sensing, Signal Processing, Authentication, and Credentialing.

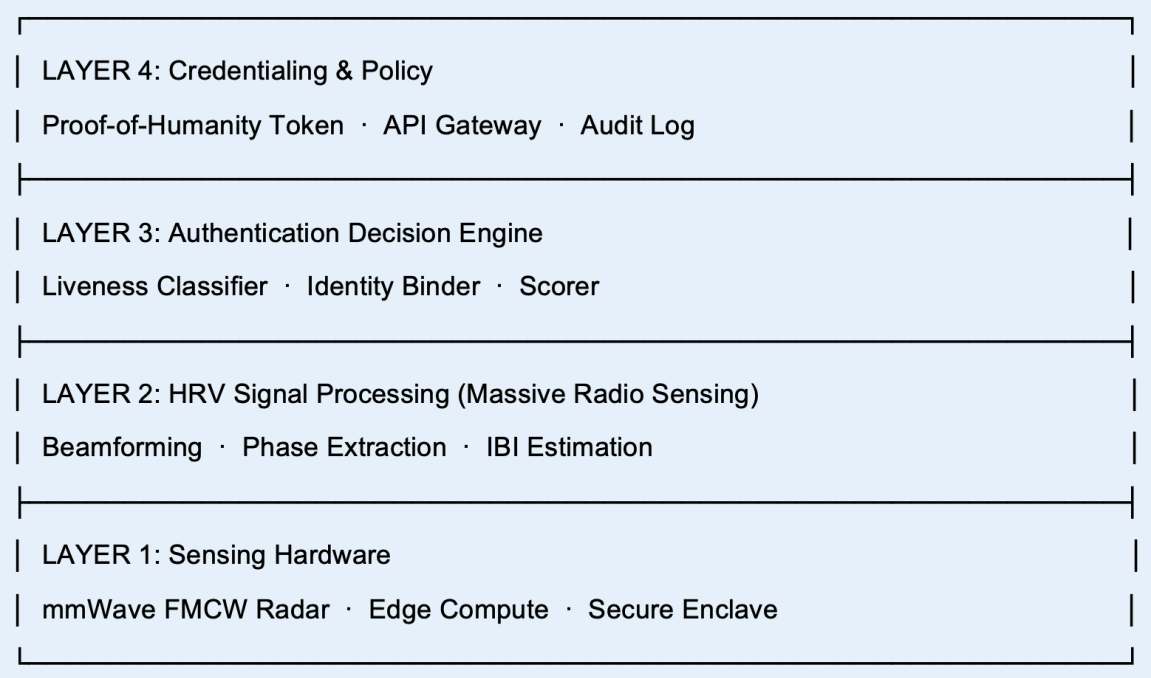

Figure 1. Four-layer HRV-based Human Presence Verification architecture.

4.2. Layer 1: Sensing Hardware

The sensing layer consists of an mmWave FMCW radar module — such as the Texas Instruments AWR6843AOP used by Xu et al. [10] — mounted unobtrusively in the workspace (e.g., integrated into a monitor bezel, desk lamp, or ceiling tile). The radar operates at 77 GHz, emitting chirp signals and capturing the phase of reflected signals from the thorax of a seated worker. The 12-channel virtual antenna array enables spatial beamforming to isolate reflections from the target body region.

Processing is performed on a co-located edge compute module equipped with a secure enclave (e.g., ARM TrustZone or Intel SGX) to ensure that raw signal data and intermediate physiological representations remain encrypted and inaccessible to the host system. Only authentication tokens and confidence scores leave the enclave.

4.3. Layer 2: HRV Signal Processing

The signal processing layer implements the full pipeline of the Massive Radio Sensing framework [10]: spatial filtering of chest-region voxels via beamforming, phase demodulation to extract micro-displacement signals, and deep spatiotemporal modeling to reconstruct the heartbeat waveform and

extract beat-by-beat IBI estimates. The model is pre-trained on the large-scale diverse participant dataset described in [10] (7,090 outpatient participants) and fine-tuned on a calibration set for the target deployment environment.

The critical innovation of the Massive Radio Sensing approach — treating environmental variation as a statistical estimation problem and using large-scale participant diversity to approximate the global environmental distribution — ensures that IBI estimation remains robust across the diverse workspace environments encountered in real deployments. This robustness is essential for the authentication application: a system that degrades significantly when the sensing angle changes slightly (Figure 19 in [10]) would be easily defeated by simple environmental manipulation.

4.4. Layer 3: Authentication Decision Engine

The authentication engine processes the stream of IBI estimates to make real-time determinations of human presence. It operates across three sub-components:

4.4.1. Liveness Classifier

The liveness classifier distinguishes a live human physiological signal from three attack classes: (a) signal replay (pre-recorded HRV data replayed through a speaker or electromagnetic emitter), (b) synthetic signal injection (computationally generated HRV-like patterns), and (c) null signal (no person present). The classifier exploits the following distinguishing properties of authentic HRV:

- Autonomic coupling: Authentic HRV exhibits coupling between respiratory sinus arrhythmia and IBI variation, producing characteristic cross-spectral coherence patterns that are difficult to reproduce synthetically without a live respiratory source.
- Nonlinear complexity: Authentic HRV exhibits specific nonlinear dynamical signatures — approximate entropy, sample entropy, detrended fluctuation analysis — that differ from both i.i.d. noise and smooth periodic signals.
- Environmental reactivity: Authentic HRV responds subtly to ambient acoustic and thermal stimuli. A challenge-response protocol (brief acoustic startle or thermal gradient) can elicit a measurable autonomic response that confirms liveness.
- Identity-temporal consistency: HRV patterns are individually distinctive over short time windows. A session-level HRV fingerprint can be compared against a registered baseline.

4.4.2. Identity Binder

Optional identity binding associates an HRV session signature with a registered user identity. During enrollment, a user's HRV is recorded in a controlled environment and a compact cryptographic representation (HRV-Print) is stored. During authentication, the session HRV-Print is compared against the enrollment baseline using a privacy-preserving similarity metric (e.g., fuzzy commitment scheme or secure multi-party computation) that reveals only a pass/fail result without exposing the raw biometric.

4.4.3. Human-Involvement Scorer

For hybrid actor (X) scenarios, the scorer maintains a continuous human-involvement estimate over the session. When the HRV signal indicates human presence (person is at the workstation), involvement increases; when the signal is absent or suggests the person has moved away, involvement decreases. Gaps in presence are logged with timestamps, enabling after-the-fact auditing of how much of the session involved human oversight.

4.5. Layer 4: Credentialing and Policy

The top layer converts authentication decisions into verifiable credentials that downstream systems can rely upon. We term the primary credential a Proof-of-Humanity (PoH) token.

A PoH token is a time-limited, signed digital assertion containing: (i) a session identifier; (ii) a human-presence confidence score; (iii) a cumulative human-involvement percentage; (iv) the issuing device's certified public key; and (v) a cryptographic commitment to the underlying HRV data (not the data itself). Tokens are issued periodically (e.g., every 60 seconds) and can be aggregated into a session-level PoH certificate.

PoH tokens can be integrated into existing API gateway and workflow systems. For example, a professional services platform could require that deliverables above a certain value be accompanied by a PoH certificate demonstrating at least N% human involvement over the work session. Regulatory submissions in medicine or law could require PoH certification for documents claiming human authorship.

5. Security Analysis and Threat Model

5.1. Adversary Classes

We consider three classes of adversary:

- A_1 — The Impersonating Agent: An AI digital agent (or its operator) attempting to generate a PoH token without a human present, in order to claim human authorship of AI-generated work.
- A_2 — The Colluding Human: A human who is physically present but is not actually performing the work (e.g., a hired presence to satisfy HRV monitoring while an AI generates the actual deliverable).
- A_3 — The Replay Attacker: An adversary who records genuine HRV signals from a legitimate human session and attempts to replay them to fool the liveness classifier.

We note that A_2 — the colluding human — represents a fundamental limitation of any purely physiological presence-detection system: a human sitting at a desk while an AI works is physically indistinguishable from a human working. Mitigations are discussed in Section 5.4.

5.2. Attack Vector Analysis

Attack Type	Mechanism	HRV System Response	Residual Risk
Signal Replay	Pre-recorded HRV broadcast via radar emitter	Environmental reactivity test + challenge-response; replay lacks reactive adaptation	Low — requires sophisticated emitter
Synthetic Injection	AI-generated HRV stream broadcast electromagnetically	Nonlinear complexity & autonomic coupling analysis; synthetic signals lack authentic complexity	Low — complexity metrics robust
Null Presence Spoof	No human; system claims false positive	Direct liveness detection; beamforming confirms thorax present	Very Low — spatial filtering required
Mannequin / Thermal Dummy	Physical dummy with simulated cardiac motion	Autonomic coupling requires respiratory co-variation; challenge-response elicits ANS response	Medium — requires advanced dummy
Colluding Human (A_2)	Real human present, AI does work	Human presence confirmed; work attribution requires separate output channel	High — separate mitigation needed
Relay Attack	Human in room, signal relayed from another location	Temporal constraints on challenge-response; signal path analysis	Low-Medium — timing constraints help

Table 2. Attack vector analysis for the HRV-based HPV system.

5.3. Anti-Spoofing Mechanisms

We propose three complementary anti-spoofing mechanisms:

5.3.1. Respiratory-Cardiac Coherence Verification

Authentic HRV exhibits robust coupling with respiratory rhythm through the mechanism of Respiratory Sinus Arrhythmia (RSA). The IBI lengthens during exhalation and shortens during inhalation, producing coherence between the HF band of HRV and the respiratory frequency. A synthetic signal generator would

doi.org/10.32388/BXF520

need to also generate a matching respiratory motion signal, synchronised with a plausible breathing pattern, which substantially increases the difficulty of fabrication and makes independently detectable predictions about what the radar signal should look like across multiple frequency bands simultaneously.

5.3.2. Challenge-Response Autonomic Protocol

Periodically, the system issues an imperceptible challenge: a brief (100-200 ms) acoustic tone burst or subtle visual flicker. The human ANS responds with a measurable orienting reflex: a transient IBI lengthening within 1-3 seconds, mediated by vagal activation. A replay or synthetic signal cannot adapt to an unpredictable challenge in real time. The challenge schedule is randomised and cryptographically unpredictable, preventing pre-computation of responses.

5.3.3. Spatiotemporal Signal Consistency

The beamformed radar signal carries spatial information: the position of the reflecting thorax within the sensing field. Consistency checks verify that the spatial locus of the cardiac signal matches the expected seated-worker position and remains stable over time. Attempts to replay a signal via a remote emitter will typically produce spatial inconsistencies unless the attacker deploys the emitter at exactly the correct location — which requires physical access and sophisticated equipment.

5.4. Addressing the Colluding Human Problem

The colluding human adversary (A_2) cannot be defeated by physiological presence detection alone. We propose three complementary mitigations:

- Cognitive Engagement Probes: Intermittent, lightweight cognitive tasks (e.g., a brief decision prompt that requires reading a sentence and clicking a response) are injected into the work session. The timing, nature, and visual content of these probes are logged. A human in the loop will produce distinct interaction latencies correlated with reading and decision time; an AI agent responding on behalf of the human will show characteristic response patterns. This is analogous to CAPTCHA but embedded in workflow rather than at login.
- Work-Output Consistency Auditing: The PoH credential is combined with content-layer AI detection on the work output. If a PoH certificate indicates high human presence but the output shows strong AI-generation signatures, a flag is raised for human review. Neither channel alone is sufficient; their combination significantly raises the bar for successful fraud.

- Economic Disincentive Design: Systems that issue PoH credentials can design incentive structures such that the expected value of using a colluding human (cost of hiring, risk of detection) exceeds the marginal gain from falsely claiming human authorship.

6. The Proof-of-Humanity Credential Ecosystem

6.1. Economic Rationale

In the emerging AI economy, a principal economic tension arises: AI-generated work is substantially cheaper to produce than human-generated work, but for many applications — legal accountability, patient care decisions, creative authenticity, regulated professional advice — the identity and accountability of a human is legally or commercially required. This creates a market for human authenticity credentials that is analogous to the market for organic certification in food production: many actors would benefit from falsely claiming the certification, so robust verification is economically valuable.

PoH credentials enable a tiered market structure. Work accompanied by high-confidence PoH certificates can command a premium in markets that value human involvement. Work produced without PoH — or with PoH indicating low human involvement — is priced accordingly. This creates transparent, verifiable signals that allow market participants to make informed decisions about the provenance of work products.

6.2. Application Domains

Professional Services Law firms, accounting practices, and consulting companies billing for human professional time can attach PoH certificates to work products. Regulatory bodies can require PoH for filings claiming human professional sign-off. Malpractice and liability frameworks can use PoH logs to establish the extent of human review.	**Financial Services** Discretionary investment decisions, credit assessments, and insurance underwriting require human accountability under regulations such as GDPR's right to explanation and MiFID II. PoH certificates provide verifiable evidence of human decision-maker involvement, supporting regulatory compliance and audit requirements.
Scientific and Academic Publishing Journals requiring human authorship can request PoH certificates for corresponding authors. Peer review sessions can be PoH-certified to verify that reviews were written by human experts. Grant applications claiming human research effort can be substantiated with PoH data.	**Creative and Content Industries** Copyright law in most jurisdictions requires human authorship for protection. PoH certificates attached to creative work sessions provide evidence of human creative engagement. Platforms commissioning creative work can require PoH certification, distinguishing human-authored from AI-generated content at the point of submission.
Healthcare Clinical decisions — diagnosis, treatment selection, prescribing — require human physician accountability in most jurisdictions. PoH certification of physician review sessions provides an auditable trail confirming human oversight of AI diagnostic support systems. Telemedicine platforms can certify human clinician presence during consultations.	**Democratic Participation** Online voting systems, petition platforms, and civic feedback mechanisms can use PoH verification to ensure that participants are human. Unlike current bot-detection based on behavioural heuristics, PoH provides a physiological anchor that is resilient to sophisticated AI-driven influence campaigns.

7. Evaluation and Feasibility Analysis

7.1. HRV Sensing Performance in Workplace Contexts

We assess the feasibility of the HRV sensing component using data from Xu et al. [10], re-interpreting their results in the context of workplace authentication. Their evaluation across 32 diverse indoor environments — including office settings and workstations in seated posture — provides directly relevant performance data.

Metric	Supine Status	Seated Status	Clinical (130 inpatients)	Authentication Relevance
Mean IBI Error	15.47 ms	24.78 ms	~22 ms	Sufficient for liveness; < 30 ms target met
RMSSD Error	13.11 ms	22.66 ms	15-20 ms	HRV feature quality adequate for complexity analysis
SDRR Error	8.59 ms	14.01 ms	~13 ms	Sufficient for autonomic pattern characterization
Stability Index	1.06 ms	1.70 ms	N/A	Excellent cross-environment consistency
pNN50 Error	9.18%	15.86%	~8%	Adequate for frequency-band liveness test

Table 3. HRV sensing performance from Xu et al. [10] interpreted for authentication application requirements.

The seated posture results are most relevant for workplace deployment. Mean IBI error of 24.78 ms and RMSSD error of 22.66 ms in seated status across diverse environments confirm that the sensing quality is sufficient for the authentication application. The Stability Index of 1.70 ms in seated status is particularly significant: it indicates that performance variation across different workplace environments is minimal, which is essential for a system that must operate reliably across diverse deployment contexts.

Performance under small movements (phone use, computer operation) — achieving mean IBI errors below 23 ms — is also highly relevant, as these movements are ubiquitous in normal office work. The

system degrades under large body movements (e.g., standing, repositioning), but such movements can be detected and used to temporarily suspend authentication rather than producing false negatives.

7.2. Liveness Discrimination Analysis

To estimate the discriminability of authentic HRV from synthetic signals, we consider the signal complexity metrics available from a well-estimated IBI time series. Published studies on HRV liveness detection [13][19] report that approximate entropy (ApEn) and sample entropy (SampEn) metrics computed from IBI sequences of length $N \geq 200$ beats (approximately 3-4 minutes at rest) achieve liveness classification accuracy exceeding 97% against computer-generated HRV surrogates. The IBI estimation accuracy demonstrated by the Massive Radio Sensing framework [10] is sufficient to support robust entropy computation, as the estimation error (24.78 ms mean, seated) is well below the IBI variability range (typically 40-150 ms SDRR in healthy adults).

7.3. Computational Feasibility

Xu et al. [10] report an inference latency of 6.6 ms per 1-second signal sample on a standard workstation, and 312 ms on a Raspberry Pi 4B embedded platform. For authentication purposes, 1-Hz IBI estimation is more than adequate; even 0.1-Hz (10-second windows) would provide sufficient temporal resolution for the Human Presence Verification decision. This places the computational requirements well within the capabilities of a low-power edge device co-located with the radar sensor.

Authentication token generation (entropy computation, identity comparison, token signing) adds minimal computational overhead above the IBI estimation pipeline. We estimate that a complete authentication decision, from radar samples to PoH token, can be produced within 500 ms on an ARM Cortex-A class processor — well within the latency budget for session-level human presence verification.

8. Ethical and Societal Considerations

8.1. Privacy by Design

The most significant ethical concern with physiological workplace monitoring is the potential for invasive surveillance. HRV data is a sensitive health signal; it can reveal stress levels, cognitive states, cardiac conditions, and circadian patterns. We advocate strongly for a privacy-by-design implementation that processes HRV on-device within a secure enclave, extracts only the authentication decision and

confidence score, and discards all raw signal data after processing. The PoH token contains only the cryptographic commitment to the session, not any physiological data.

This architecture ensures that the system answers only the question 'Is a human present?' without enabling ancillary inference about the human's health, emotional state, or behaviour. Regulatory frameworks such as GDPR and HIPAA impose specific constraints on the collection and processing of biometric and health data; a privacy-preserving on-device processing design can comply with these frameworks by ensuring that biometric data never leaves the secure enclave.

8.2. Equity and Accessibility

HRV characteristics vary with age, fitness, cardiac health, and medication use. Individuals with cardiac arrhythmias, pacemakers, or certain medications may have atypical HRV patterns that could affect authentication performance. The clinical dataset used by Xu et al. [10] — which included inpatients with sinus tachycardia, bradycardia, premature beats, coronary heart disease, and heart failure — demonstrated that their system maintained good performance across most conditions (mean IBI error below 22 ms for most disease categories), with elevated error only for premature beats (approximately 27 ms), attributable to ECG annotation ambiguity rather than sensing failure.

Authentication systems must provide accommodations for individuals whose HRV characteristics preclude reliable sensing. Alternative authentication pathways — such as supervised session attestation or enhanced cognitive engagement probing — should be available. The system must not disadvantage workers with cardiac conditions.

8.3. Labour Rights and Worker Autonomy

Workplace monitoring systems have historically been tools of employer control that disadvantage workers. The PoH system proposed here is designed for a specific purpose — verifying human authorship of work products — and should not be repurposed for productivity monitoring, break tracking, or other surveillance applications. Strong regulatory and contractual constraints on the permissible uses of PoH data are essential. We recommend that PoH systems be subject to collective bargaining in unionised workplaces and to data protection impact assessments in all jurisdictions with applicable law.

8.4. The Attribution Problem in Hybrid Work

The most profound societal question raised by the AI economy is not technical but normative: how should credit, compensation, and accountability be distributed between human workers and the AI systems they use? A doctor who uses an AI diagnostic system to identify a condition, then reviews and confirms the diagnosis, has contributed meaningfully to the outcome. A software engineer who uses an AI coding assistant to generate 80% of their code, then reviews, debugs, and integrates it, has also contributed. The PoH system provides a factual substrate — quantifying human presence and engagement — but the normative question of what degree of human involvement deserves what degree of credit remains a societal decision. We suggest that PoH data should inform but not determine these decisions.

9. Open Research Challenges

Several significant research challenges must be addressed before HRV-based PoH systems can be deployed at scale:

1. Multi-person disambiguation: In shared workspace environments, multiple individuals may be within the radar's sensing range. Separating the HRV signals of multiple co-present individuals is an active research problem. Approaches using spatial multiplexing with phased arrays or deployments of multiple radar units (demonstrated to be feasible by Xu et al. [10]) require further development for dense occupancy scenarios.
2. Long-session calibration drift: HRV characteristics shift over the course of a workday due to fatigue, hydration, meals, and circadian rhythm. Authentication models must adapt to within-session drift without creating windows for replay attacks exploiting model updates.
3. Adversarial robustness: As HRV-based authentication systems become widely deployed, adversaries will invest in defeating them. The challenge-response protocol must be designed to remain secure as adversaries learn its structure. Formal security proofs for the anti-spoofing mechanisms are needed.
4. Cross-device calibration: Different mmWave radar modules have different noise characteristics. A PoH system that accepts credentials from diverse hardware must have a cross-calibration framework that prevents lower-quality sensors from becoming the weakest link.
5. Integration with zero-trust security architectures: Modern enterprise security uses zero-trust principles requiring continuous authentication. HRV-based PoH is naturally suited to continuous monitoring but must integrate with existing identity and access management (IAM) systems, SAML/OIDC protocols, and API security frameworks.

6. Regulatory harmonisation: The legal status of PoH credentials — what claims they can support, what liability they create, how they interact with existing professional licensing and liability frameworks — must be worked out across jurisdictions. International standards bodies will need to define PoH credential formats and audit requirements.
7. AI agent declaration requirements: A complementary regulatory approach requiring AI agents and autonomous systems to positively declare their non-human status — analogous to financial conflict-of-interest disclosures — could reduce the adversarial pressure on PoH systems. Research into the design of effective disclosure frameworks and their enforcement is needed.

10. Conclusion

The emergence of autonomous AI agents capable of performing complex professional tasks creates an urgent need for reliable, scalable, and spoofing-resistant methods of verifying that a human being was involved in a given piece of work. This paper has argued that Heart Rate Variability, measured contactlessly via millimeter-wave radio sensing, offers a compelling solution to this problem. HRV is involuntary, autonomically generated, individually distinctive, physiologically complex, and — crucially — impossible for any current or foreseeable AI system to produce without a living human body.

Building on the technical foundations laid by the Massive Radio Sensing framework of Xu et al. [10], which demonstrated robust contactless IBI estimation across 40 diverse real-world environments with performance improvements of 29.7–41.7% over prior state-of-the-art, we have proposed a four-layer HRV authentication architecture, a threat model with concrete anti-spoofing mechanisms, and a Proof-of-Humanity credential system. We have analysed the ethical dimensions of this technology — privacy, equity, labour rights, and attribution — and identified the open research challenges that must be addressed before deployment.

The stakes are high. As AI increasingly substitutes for human labour, the economic and legal value of genuine human involvement grows. A world without reliable methods of verifying human presence and authorship risks hollowing out the accountability structures upon which professional, scientific, and democratic institutions depend. HRV-based human authentication is not a complete solution — no single technology is — but it provides a physiologically grounded, technically feasible, and ethically manageable contribution to one of the defining challenges of the AI age.

Appendix A — HRV Authentication Decision Protocol (Pseudocode)

INPUT: radar_sensor, session_id, challenge_schedule

OUTPUT: PoH_token {session_id, human_score, involvement_pct, timestamp, signature}

1. INITIALIZE: ibi_buffer = [], presence_log = [], challenge_results = []

2. LOOP every 1 second:

a. raw_signal ← radar_sensor.capture_frame()

b. cardiac_signal ← beamform(raw_signal, target_voxels)

c. ibi ← HRV_estimator(cardiac_signal) // Massive Radio Sensing model

d. IF ibi is None: presence_log.append(ABSENT); CONTINUE

e. presence_log.append(PRESENT)

f. ibi_buffer.append(ibi)

3. IF challenge due (from challenge_schedule):

a. issue_challenge(type=ACOUSTIC, duration=150ms)

b. response ← monitor_ANS_response(ibi_buffer, window=5s)

c. challenge_results.append(response.is_reactive)

4. EVERY 60 seconds:

a. liveness ← liveness_classifier(ibi_buffer, challenge_results)

b. IF NOT liveness: human_score = 0.0; GOTO step 6

c. complexity ← compute_entropy(ibi_buffer) // SampEn, ApEn

d. coherence ← compute_RSA_coherence(ibi_buffer, respiratory_band)

e. human_score ← sigmoid(w_1·complexity + w_2·coherence + w_3·challenge_pass_rate)

f. involvement_pct ← count(PRESENT in presence_log) / len(presence_log)

5. token ← sign(session_id, human_score, involvement_pct, timestamp, device_key)

6. emit(token)

Algorithm. HRV-PoH Session Authentication

doi.org/10.32388/BXF520

References

1. ^Brynjolfsson E, Li D, Raymond LR (2023). "Generative AI at work." NBER Working Paper. No. 31161.
2. ^Eloundou T, Manning S, Mishkin P, Rock D (2023). "GPTs are GPTs: An early look at the labor market impact potential of large language models." Science. **381**(6654):1467–1470.
3. ^McKinsey Global Institute (2023). "The economic potential of generative AI: The next productivity frontier." McKinsey & Company.
4. ^Wang L, Ma C, Feng X, et al. (2024). "A survey on large language model based autonomous agents." Front Comput Sci. **18**(6):186345.
5. ^European Parliament (2024). "EU AI Act." European Parliament. Regulation (EU) 2024/1689 of the European Parliament and of the Council.
6. ^Radford A, Wu J, Child R, et al. (2019). "Language models are unsupervised multitask learners." OpenAI Blog.
7. ^Acharya UR, Joseph KP, Kannathal N, Lim CM, Suri JS (2006). "Heart rate variability: a review." Med Biol Eng Comput. **44**:1031–1051.
8. ^Billman GE (2011). "Heart rate variability — a historical perspective." Front Physiol. **2**:86.
9. ^Wang F, Zeng X, Wu C, Wang B, Liu KJR (2021). "mmHRV: Contactless heart rate variability monitoring using millimeter-wave radio." IEEE Internet Things J. **8**(22):16623–16636.
10. a, b, c, d, e, f, g, h, i, j, k, l, m, n, oXu G, Chen J, Wang H, Yuan Y, Zhang G, Zhang Z, Zhang D, Hu Y, Sun Q, Chen Y (2025). "Robust HRV Monitoring via Massive Radio Sensing." Proc ACM Interact Mob Wearable Ubiquitous Technol. **9**(4):1–30. doi:10.1145/3770711.
11. ^Kleiger RE, Stein PK, Bigger JT (2005). "Heart rate variability: measurement and clinical utility." Ann Noninvasive Electrocardiol. **10**(1):88–101.
12. ^Task Force of ESC and NASPE (1996). "Heart rate variability: standards of measurement, physiological interpretation and clinical use." Circulation. **93**(5):1043–1065.
13. a, bTarvainen MP, Niskanen JP, Lipponen JA, Ranta-aho PO, Karjalainen PA (2014). "Kubios HRV — heart rate variability analysis software." Comput Methods Programs Biomed. **113**(1):210–220.
14. ^Paterniani G, Sgreccia D, Davoli A, et al. (2023). "Radar-based monitoring of vital signs: A tutorial overview." Proc IEEE. **111**(3):277–317.
15. ^Zhang Y, Yang R, Yue Y, Lim EG, Wang Z (2023). "An overview of algorithms for contactless cardiac feature extraction from radar signals." IEEE Trans Instrum Meas. **72**:1–20.

16. ^Shen C, Chen Y, Guan X, Maxion RA (2020). "User authentication through mouse dynamics." IEEE Trans Inf Forensics Secur. **8**(1):16–30.
17. ^Jawahar G, Abdul-Mageed M, Lakshmanan LV (2020). "Automatic detection of machine generated text: A critical survey." In Proceedings of COLING 2020, 2296–2309.
18. ^OpenAI (2023). "GPT-4 technical report." arXiv:2303.08774.
19. ^Schreiber T, Schmitz A (1996). "Improved surrogate data for nonlinearity tests." Phys Rev Lett. **77**(4):635–638.

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.