Position: Olfaction Standardization is Essential for the Advancement of Embodied Artificial Intelligence

Anonymous Author(s)
Affiliation
Address
email

Abstract

Despite extraordinary progress in artificial intelligence (AI), modern systems remain incomplete representations of human cognition. Vision, audition, and language have received disproportionate attention due to well-defined benchmarks, standardized datasets, and consensus-driven scientific foundations. In contrast, olfaction—a high-bandwidth, evolutionarily critical sense—has been largely overlooked. This omission presents a foundational gap in the construction of truly embodied and ethically aligned super-human intelligence. We argue that the exclusion of olfactory perception from AI architectures is not due to irrelevance but to structural challenges: unresolved scientific theories of smell, heterogeneous sensor technologies, lack of standardized olfactory datasets, absence of AI-oriented benchmarks, and difficulty in evaluating sub-perceptual signal processing. These obstacles have hindered the development of machine olfaction despite its tight coupling with memory, emotion, and contextual reasoning in biological systems. In this position paper, we assert that meaningful progress toward general and embodied intelligence requires serious investment in olfactory research by the AI community. We call for cross-disciplinary collaboration—spanning neuroscience, robotics, machine learning, and ethics—to formalize olfactory benchmarks, develop multimodal datasets, and define the sensory capabilities necessary for machines to understand, navigate, and act within human environments. Recognizing olfaction as a core modality is essential not only for scientific completeness, but for building AI systems that are ethically grounded in the full scope of the human experience.

1 Introduction

Over the past two decades, artificial intelligence (AI) and machine learning (ML) have undergone rapid development, steadily advancing toward the vision of super-human intelligence. This progress has been underpinned by the creation of objective, measurable benchmarks that enable rigorous model evaluation and reproducibility. AI systems now equate or surpass human performance in tasks ranging from medical diagnostics [123, 80] and image synthesis [121] to speech recognition [114] and fluent language communication [8]. In embodied intelligence, advances in tactile sensing allow robots to perceive stimuli with sensitivities that exceed human thresholds. Roboticists, in parallel, are building humanoid platforms capable of replicating everything from human gait [65, 115, 69] to dexterous hand manipulation [137, 45, 10], and even abstract reasoning capabilities [29, 43, 152].

Yet, among the senses that define human embodiment, two remain largely absent in artificial systems: taste and smell. While the exclusion of gustation may be pragmatically justified—given its primary role in nutrient acquisition in humans—olfaction is far more integral to human cognition. As the third-highest bandwidth sensory modality after vision and audition, olfaction is uniquely linked to memory, emotional response, and decision-making. Its exclusion raises an important question: can we

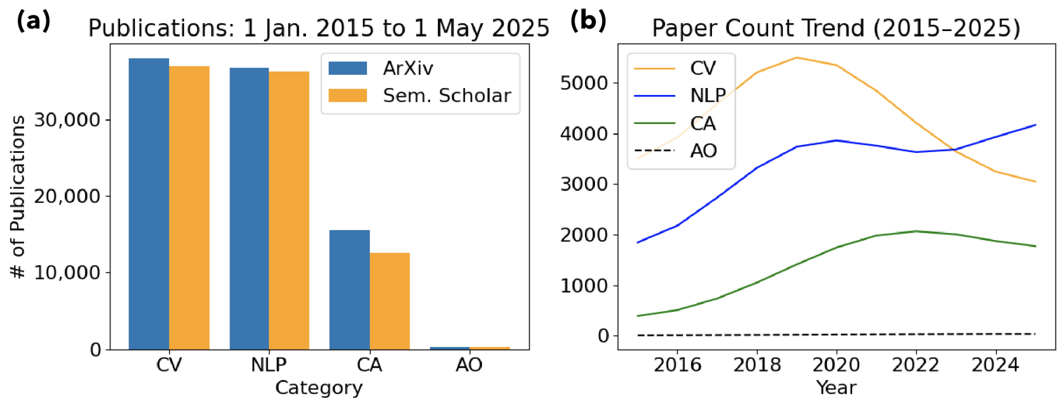

Figure 1: The number of publications from ArXiV [38] and Semantic Scholar [86] over the period 1 January 2015 to 1 May 2025 in computer vision (CV), natural language processing (NLP), computer audio (CA), and artificial olfaction (AO). **(a)** illustrates the total number of publications in each category as retrieved from the ArXiv and Semantic Scholar databases. **(b)** shows the temporal trend in published research for each category, with olfaction remaining stagnant over the last ten years.

claim to be in pursuit of building artificial general intelligence if we ignore one of the core modalities through which humans and other animals perceive, interpret and navigate in their environments?

Despite its clear relevance, olfaction has been disproportionately neglected within the AI community [39, 85, 32]. The number of papers published in machine olfaction was less than 1% of the volume of computer vision and natural language processing over the last decade (see Figure 1 and Section A). This omission not only limits the completeness of embodied AI but also overlooks a rich domain of cognitive science that remains underexplored. We argue that this neglect stems from several compounding challenges—challenges that, while substantial, are tractable. Specifically, we identify five interconnected systemic gaps that can be addressed through a concerted effort by the AI community:

1. **Scientific Understanding of Olfaction:** There is no unified scientific consensus on the underlying **mechanisms of olfaction**, with multiple competing hypotheses still under debate.
2. **Olfaction Data Standard:** Olfactory data can be captured via heterogeneous sensor modalities, leading to a lack of **standardized data representation** and hardware specification.
3. **Objective Annotation:** Sensor measurements often detect signals below the threshold of human observability, thus complicating **objective evaluation** and ground-truth labeling.
4. **Olfaction Datasets:** AI lacks peer-reviewed, large-scale **olfactory datasets**, particularly those that are multimodal or suitable for training modern machine learning systems.
5. **Olfaction Benchmarks:** The lack of established AI-specific **benchmarks** inhibits communal advancement of olfactory perception, generation, reasoning, and sensor hardware.

Our Position: We contend that the slow progress and inequity in machine olfaction—especially within artificial intelligence—can be attributed to these five systemic gaps. Addressing even a subset of these issues, whether through foundational research, infrastructure investment, or cross-disciplinary collaboration, could catalyze rapid advancements in the field and advance the state of the art of artificial olfaction equitable to vision, audition, and language to enable another facet of embodied AI.

We therefore call for a focused research agenda, increased funding, and dedicated talent toward machine olfaction—recognizing it as a critical enabler for truly embodied artificial intelligence.

2 The Case for Olfaction

2.1 Scientific Understanding of Olfaction

Olfaction is a unique sense in that multiple modalities can be used to detect an odourant. Some sensors "see" odourants by measuring the change in optical wavelength bands as a molecule passes. Other detectors measure small chemical reactions that occur through electron, oxygen, or other molecule transfer across a diffusion medium, such as an electrolyte. Still other sensors "hear" odourants by measuring the vibrational modes of the compound at a quantum level. The first two exemplify what is commonly referred to as the Shape Theory of Olfaction (STO) where a molecule's physical properties are hypothesized as the most significant contributors to its aroma [147] [16] [129]; the latter alludes to what is now commonly referred to as the "Vibrational Theory of Olfaction" (VTO).

First postulated by Dyson in 1928 [55] and again supported in 1938 [98], the vibrational theory of olfaction was originally dismissed as experimental data via Raman spectroscopy was gathered as evidence against the idea. Luca Turin re-popularized the idea in 2001 [139] claiming that Dyson's proposal was not incorrect, but perhaps slightly uncalibrated. Turin suggested that the detection mechanism in mammalian olfactory receptors is due to inelastic electron tunneling. Brooks, et al. discuss the feasibility of a derivative of VTO called phonon-assisted tunneling in [22]. Block, et al. have strongly refuted the plausibility of VTO in [17], leaving an unsettled debate on the matter.

Both VTO and STO provide convincing evidence on why each theory is plausible, and both sides have growing experimental data to support. "Neither theory can fully explain why the scent of some molecules are concentration dependent. This is a problem well known to perfumers, and yet unexplained" [20]. Stevens reports in his 1951 work that "it is very probable that no one physical property alone is involved in the physical nature of the adequate stimulus" [133]. This leaves the science of olfactory sensing still open with no single sensor technology acting as a unified theory.

How, then, does one create a data standard for a modality that does not yet have a consensus on its scientific functionality? From our position, this dichotomy in understanding is not a barrier to entry for defining olfactory datasets, but an opportunity for the AI community. The JPEG and PNG standards did not require a deep understanding of how images are interpreted by the human visual cortex (although we admit that understanding the biology of vision has influenced the larger progress of computer vision). Aircraft do not replicate nature's form of flight, yet there are standards that each airplane must pass to prove airworthiness for operation. In this manner, we believe that it is important to keep the plausible theories of olfaction in tight consideration whilst developing a corresponding standard, but the scientific progress and standardization process can move forward in parallel.

2.2 Olfaction Data Standard

In contrast to other sensory modalities, olfaction lacks a universally accepted data standard. In vision, for instance, most visible colours can be approximated using combinations of red, green, and blue (RGB) channels, reflecting the trichromatic nature of human colour perception. As a result, digital images are typically discretised using this encoding scheme. Visual information is typically stored in well-defined formats such as PNG or JPEG (for images) and MP4 or AVI (for videos), which facilitate uniform processing and widespread data sharing. Similarly, audio files are binned according to different frequencies. Audio data benefits from standardized representations like WAV files, while speech can be transcribed into words (or tokens) that serve as a natural language standard. These common formats in each modality have been instrumental in driving rapid progress in AI systems.

Why does olfactory data not have such analogue?

Sensor Modality. Canines can detect and navigate to odours down to concentrations in the parts-per-trillion (ppt) regime [36]. For humans, the odour detection threshold in air ranges between sub-parts-per-billion (ppb) to hundreds of parts-per-million (ppm) [90]. Such low detection thresholds are achieved by pooling and averaging millions of receptor neurons across the nasal epithelium [1], each of them with a high sensitivity to a particular molecule or group of molecules.

While computer vision and audition models operate with data rates and compression schemes comparable to human input channels, there is still no consensus in machine olfaction on the optimal sensor modality. Deploying and interfacing biological olfactory receptors may be tempting, however,maintaining their viability over extended periods remains technically challenging, inevitably leading to stability issues [130]. As a result, a range of alternative sensing elements are employed, including electrochemical sensors, conducting polymer composite gas sensors, metal-oxide (MOx) gas sensors, optical gas sensors, acoustic sensors, and carbon nanotube-based sensors. These sensor types differ in their performance characteristics—such as sensitivity, selectivity, and response time—as well as in their operational constraints, including power consumption, physical footprint, and long-term stability (e.g. susceptibility to drift). Sensor selection is therefore typically application-specific.

Bandwidth. Recent work by Zheng and Meister [157] quantifies a long-suspected limitation of human cognition: humans' total sensory input is ingested at $\approx$1 GB/second. However, despite the brain's massive internal bandwidth, our expressive output saturates at approximately 10 bits per second. This is consistent with prior estimates of speech production rates and reflects an inherent dichotomy between high-bandwidth perception and low-bandwidth action/communication channels. For AI researchers and sensory neuroscientists, this bottleneck reframes the importance of input channels in artificial systems. As AI becomes embodied and more exploratory, there is a growing need to diversify and optimize sensory bandwidth and compute beyond vision, language, and audition.

For humans, vision is the highest bandwidth modality. It is estimated that the afferent visual interpretation is at a rate of 20 MB/sec [157]. On a smaller scale, the human auditory system processes 150-200 KB/s (although only 12 KB/s may be perceived). What, however, constitutes the data ingestion rate of human olfaction? We can compute this from some first-order calculations.

For human olfaction, we approximate data rate based on the following assumptions:

- Sniffing rate of $\approx$1 Hz
- 400 olfactory receptor neuron (ORN) types, converging onto two glomeruli each [107]
- 25 mitral cells for each glomerulus [71]
- 0 - 13 spikes in each mitral cell per sniff cycle [51] (minimally represented by 4 bit)

Rolling these calculations out yields >5 kB/sec for the human olfactory code, making olfaction our third-highest throughput sense. Canines, on the other hand, see the world through their nose: With approximately 2.5× receptor diversity, and an up to 8× higher sniffing rate [122], they likely perceive olfactory data at >100 kB/s. This 20× performance over humans is clear evidence to support why superhuman olfaction is achievable in multimodal and embodied AI.

In machine olfaction, slow sensor dynamics have historically limited real-time use. Only recently have advances in hardware and software culminated towards closing this gap, where for instance an electronic nose with millisecond-scale response times [46] was demonstrated to outperform mice in temporal discrimination tasks. Such developments bring artificial olfactory bandwidth closer to biological levels, enabling new real-time applications that previously constrained development.

Encoding. Both vision and audition process stimuli that vary continuously in physical properties (frequency of light and air waves, respectively), with corresponding continuous receptor mappings and perceptual representations. In contrast, olfaction forms a discrete space: As each odourant molecule has a unique molecular structure, there is no inherent analogous to the frequency. The vast diversity of odourant molecules simply does not align along a single, continuous dimension.

Furthermore, each ORN can bind to multiple odourant molecules, and each odourant can bind to multiple ORNs, resulting in a combinatorial coding scheme [99]. At high odour concentrations—typically in the ppm range or above—odour representation at the receptor and glomerular levels relies on the broad tuning properties of olfactory receptors, leading to overlapping activation patterns across many receptor types [99, 96, 105]. However, such high concentrations are uncommon in natural environments, where odourant levels rarely exceed the ppb range [141]. Under these more ecologically relevant conditions, olfactory sensory neurons in mammals exhibit far sparser activation, with some receptors responding selectively to just a single odourant at low concentrations [26, 37, 48].

Olfactory signals are not just encoded sparsely in receptor activation, but also in time [48]. In natural environments, airborne odours are carried in turbulent plumes, which are highly intermittent in both space and time [27]. Rather than forming a continuous stream, odour molecules arrive at a sensor or nose in brief, irregular bursts, separated by periods in which no odour is detected. These gaps can vary greatly in duration, depending on factors such as wind direction, turbulence, and distance from the source. As a result, animals often encounter odours in a fragmented and unpredictable manner, sometimes experiencing milliseconds-long bursts, followed by seconds or more of clean air [27, 154]. The brain's capacity to extract meaningful information from these fleeting encounters suggests that olfactory systems must be tuned for episodic, rather than continuous, sampling of the chemical world.

Towards an olfactory data standard. The physics of olfaction suggest that we must think about it differently than its counterparts. Much like proprioception or vestibular inputs in biology, olfaction informs action and decision-making without requiring high-bandwidth symbolic output [157]. These properties lend olfaction well to neuromorphic computing, a young counterpart to the Von Neumann architecture [120]. Neuromorphic computing is designed to be more energy efficient, highly parallel, and facilitate event-based processing versus Von Neumann clock-based processing [116, 117, 88]. Progress toward defining an olfactory data standard is not contingent on the selection of the processor, but the properties of olfaction are inherently optimized by the neuromorphic architecture. AI will naturally require more power, data, and compute as it scales and embodies more capabilities

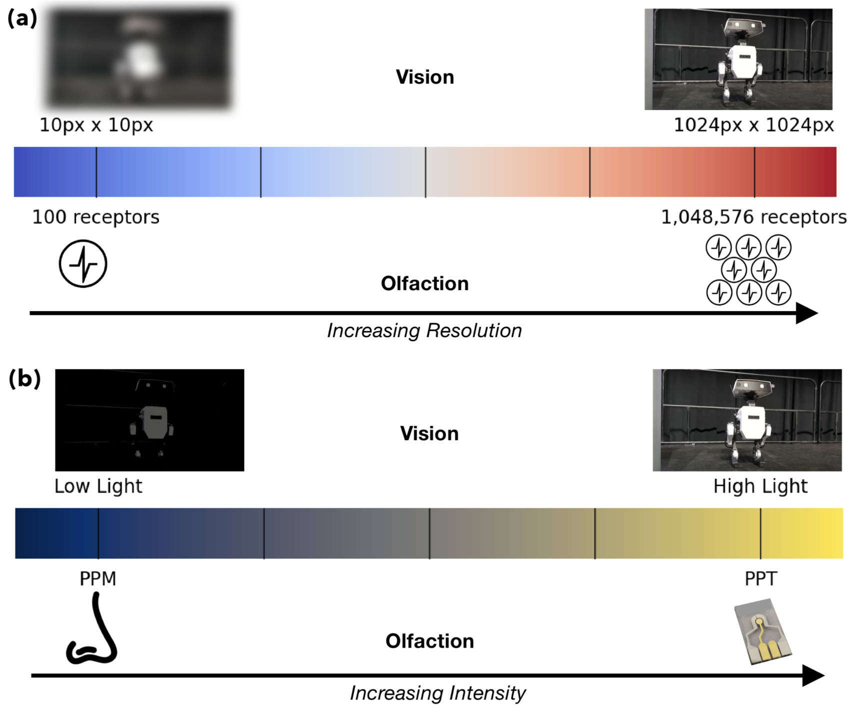

Figure 2: Analogy of olfaction to computer vision. **(a)** The number of receptors in olfaction can be likened to pixel count in an image, or its *resolution*. **(b)** The *intensity* of an olfaction sensor is synonymous with the dynamic range in a camera. In other words, the number of molecules that dock to the olfaction detector is analogous to the number of photons captured by the camera detector.

[2, 82]. Consequently, it would be prudent to keep in mind alternative architectures in the pursuit of standardizing the sense of smell. Addressing the opportunities we delineate here could be the catalyst for research into new computational architectures driven by olfaction. **As a result, we believe that AI researchers and practitioners defining a data standard for olfaction is the first call to action.**

2.3 Objectivity

Human olfaction is inherently subjective. For instance, a stroll through a village in France may evoke delight from the aroma of freshly baked baguettes or crêpes—described by humans as "fresh," "sweet," or "roasted." These descriptors are inherently linguistic, culturally influenced, and shaped by personal experience. In contrast, olfactory sensors interpret scent via molecular signatures: discrete patterns of volatile organic compounds (VOCs) that can be digitized and processed by machines. While sensors may detect compounds like *2-acetyl-1-pyrroline*, the molecule responsible for the smell of baked bread, humans do not perceive this molecular identity; we perceive semantic impressions.

This divergence between molecular and semantic perception introduces a profound challenge: how do we construct objective olfactory benchmarks when the human experience of smell is so context-dependent? The mapping from olfactory receptor activation to semantic descriptors is not deterministic. It is modulated by individual genetic variation [148], environmental context [97], health status [21, 91], age [53], and even neurocognitive conditions [52]. Concepts like the *Principal Odour Map* from Lee, et al. [89] highlight this, noting non-unanimous descriptors for the same aroma, and research on baked bread [93] shows varied descriptors due to differing emitted compounds. Figure 4 shows the limits of human perceptibility and sensor detectability for molecular associated with the aroma of baked bread. This subjectivity directly impacts our ability to create a unified, multimodal concept space (refer Figure 3). Without a common objective and grounding, olfactory data cannot be meaningfully integrated with other modalities to develop enhanced multi-modal AI systems.

Such biases, however, are not unique to olfaction. In vision, color perception varies across individuals—some are red-green color blind, while others perceive ultraviolet spectra. Yet, despite these variances, the scientific community has converged on objective representations. The color "orange" is mapped to specific wavelengths of visible light, and though linguistic interpretations may vary, consensus exists about the underlying physical measurements. Computers represent *orange*

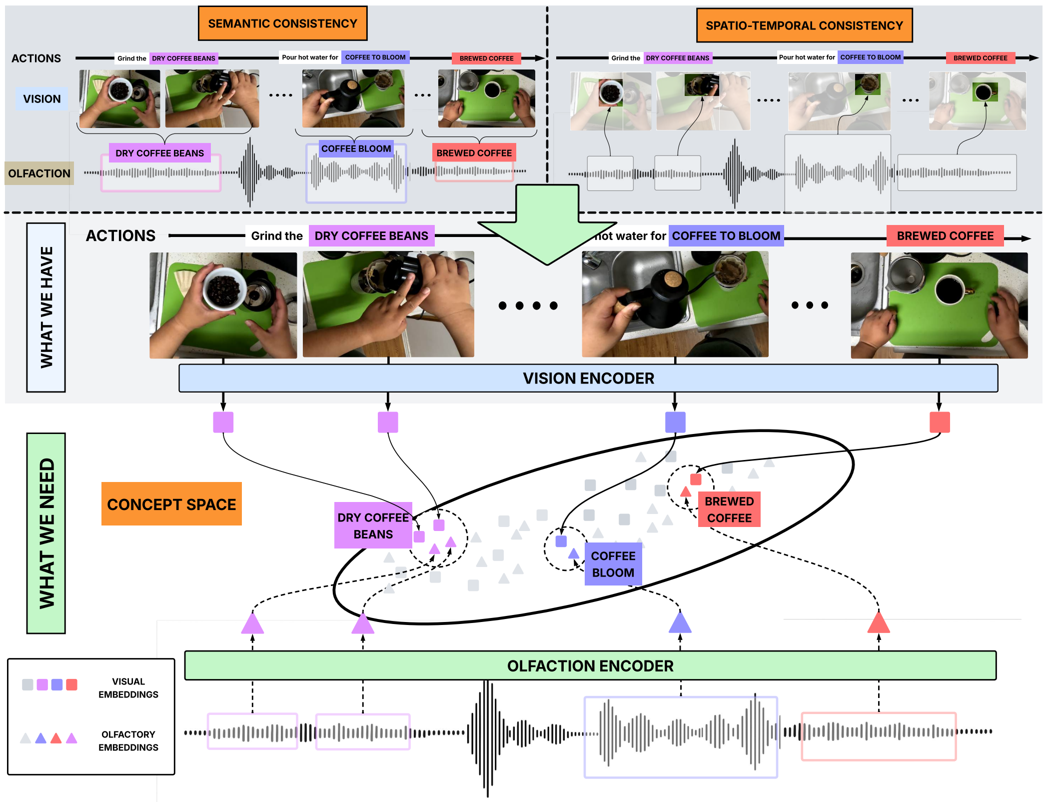

Figure 3: Illustrates how standardized olfactory data and models can integrate with other modalities (e.g., vision) for learning enhanced *Large Multi-modal Concept Models*. **Top:** We highlight two key properties: **semantic consistency**, where, for example, the visual presence of dry coffee beans, coffee bloom, and brewed coffee is consistently paired with their respective olfactory signatures. It also shows **spatio-temporal consistency**; while both visual and olfactory data changes are conceptually synchronous (e.g., the sight and smell of brewing coffee occur together), they are captured at different frequencies or temporal resolutions. **Bottom:** We contrast current, siloed approaches *what we have* with the goal of future systems *what we need*. The envisioned future state, enabled by the standardization efforts advocated in this paper, shows olfactory embeddings contributing to a more comprehensive and unified concept space leading to richer understanding of the real-world.

as quantifiable spectral data—e.g., 590–620 nm—independent of individual perception. The more abstract visual patterns they identify are ultimately represented as vectors in model latent space [64]. A similar paradigm is essential for olfaction if we are to develop the large-scale, standardized datasets (see Section 2.4) required for modern machine learning. Projects like *GoodScents* [35] and *Leffingwell* [9] attempt to map molecules and human odour descriptors but these are limited. To move towards the robust datasets envisioned, we must prioritize the collection of raw, digitized sensor data capturing objective molecular signatures rather than relying on often inconsistent human labels.

The ability of systems to operate beyond human perceptual limits is key in olfaction, much like in computer vision. For instance, while an odourant at 2 ppb may be undetectable by humans, it is often fully observable by olfactory sensors or canines. This is comparable to how an object at low visual resolution (e.g., 10×10 pixels, see Figure 2) becomes clearer with increased resolution. Consequently, just as vision models can learn abstract features from low-resolution inputs, olfactory models could identify molecular patterns at these sub-perceptual concentrations. To harness and measure this potential, establishing objective measures and verifiable ground truth is a prerequisite for robust benchmarking of olfactory progress (see Section 2.5).

2.4 Olfaction Datasets

The quest for truly Embodied AI cannot succeed while olfaction, a primary sensory modality, remains in the digital shadows. Can we, as a community, afford to delay the creation of an olfactory equivalent to ImageNet [44]? The current lack of consensus on digital odour representation and the consequent dearth of large-scale, standardized datasets are not mere inconveniences; they are fundamental roadblocks. This void actively stifles progress in developing machine learning models capable of sophisticated odour perception and prediction—capabilities that are paramount for robots and AI systems designed to operate in, understand, and interact with the complexities of the real world.

Pioneering efforts like the Principal Odour Map (POM) [89] and Eigengraphs [135] leverage on existing databases such as *GoodScent* [35], *LeffingWell* [9] and *DoOR* [104]. While these offer glimpses into "digitizing" odours, they also underscore critical challenges, namely the profound subjectivity of human aroma classification (see Section 2.3) and the limited scope of current datasets [126]. From a machine learning perspective, two datasets are particularly noteworthy and frequently used. One is the psychophysical dataset by Keller et al. [84], which links 480 diverse molecules to human perceptual ratings across intensity, pleasantness, and semantic descriptors. It highlights reliable structure–percept correlations and underscores the strong influence of familiarity on verbal reports, making it a key resource for modelling olfactory perception. The other is a large-scale gas sensor dataset from Vergara et al. [140], capturing responses from 72 MOx sensors exposed to 10 gases under turbulent flow in a custom wind tunnel. Designed to mimic open-air sampling conditions, it includes 18,000 time series measurements and supports algorithm development for robust odour detection in complex, variable environments. However, later work revealed that certain limitations of the experimental protocol during data collection may have compromised the dataset's reliability [47] and thus lead to overly optimistic algorithm evaluation results [49]. Thus the current fragmentation in olfactory data representations presents a significant barrier to data aggregation.

The assertion that we have reached "peak data" [136] overlooks the vast, uncharted territory of olfactory information. This is not merely about adding another modality; it is about unlocking a fundamentally new dimension of data for multimodal AI, pushing it beyond the confines of current internet-trained models. We contend that developing a robust olfactory data standard, in parallel with the construction of large-scale datasets (much like audio research progressed with varied representations before full convergence [30]), is essential for the next leap in multi-modal AI (see Figure 3), leading to omni-perceptual embodied and potent systems.

For Embodied AI to advance beyond its current sensory limitations, we must prioritize the development of olfactory datasets that capture the richness of real-world olfactory experiences in a structured, machine-learnable format. This necessitates a paradigm shift towards datasets that encompass: (a) **Multimodal Datasets for Static Scene Understanding:** Olfaction does not operate in a vacuum. To build AI systems that achieve a holistic understanding of their surroundings, even in relatively static contexts [44, 92, 41], we require datasets that integrate olfactory information with static 2D/3D scenes. (b) **Spatio-Temporal Olfactory Archives for Dynamic Scene Understanding:** For embodied agents to effectively operate and navigate within dynamic environments, and to understand complex, unfolding activities, a different class of olfactory datasets is essential. These datasets must map the olfactory world in four dimensions similar in-spirit to multi-modal 4D datasets [66, 110, 57], comprising time-series olfactory data captured by mobile sensor arrays. Critically, this data needs to be meticulously correlated with the agent's trajectories, the geometry of the environment, and ground-truth locations of odour sources. Such spatio-temporal olfactory archives are indispensable for training AI in crucial tasks like hazardous leak detection, olfactory search-and-rescue operations, and for enabling a deeper understanding of activities that leave olfactory traces over time and space.

To overcome the inherent variability of olfactory perception and ensure the utility of these datasets, our position is that: (1) The community should converge on robust methodologies by prioritizing the collection of raw, digitized sensor data under rigorously controlled environmental conditions to capture objective molecular signatures. Experimental controls must ensure that odourant concentrations remain within detectable thresholds for most humans. It has been shown [150] that the majority of biologically relevant odourants occur below the parts-per-billion (ppb) concentration range, necessitating highly sensitive and reproducible instrumentation. (2) We must develop and adopt standardized protocols for sensor calibration, data acquisition, and comprehensive annotation akin to recent efforts from robotics community [34]. (3) When incorporating subjective human descriptors, it's crucial to implement systematic approaches that utilize multiple annotators and consensus-scoring to mitigate individual bias and address concentration-dependent perceptual variations [87, 125, 83].

2.5 Benchmarks: Charting the Course for Olfactory-Empowered Embodied AI

Without robust and standardized benchmarks, the development of olfactory capabilities within Embodied AI will remain disjointed and progress unmeasurable. It is imperative that we, as a research community, define a suite of benchmark tasks that not only test olfactory perception in isolation but, more critically, evaluate its contribution to holistic understanding of dynamic scenes.

These benchmarks, drawing inspiration from successes in Computer Vision (e.g. ImageNet [44], COCO [92]) and NLP (e.g. GLUE [142], SQuAD [119]), must provide clearly defined tasks, curated multi-modal datasets (as discussed in Section 2.4), and rigorous evaluation metrics. We propose a focus on the following benchmarks, where olfactory data is not just an add-on, but a **transformative component** in understanding and reasoning:

Foundational Olfactory Perception. This paradigm is centered on the basic ability of machine learning systems to sense olfactory stimuli. It encompasses recognition tasks such as odour component identification which focus on determining the presence and concentrations of specific volatile organic compounds. In principle, this is analogous to established tasks like image classification [44] or audio event detection [124, 111]. Progress within this paradigm is exemplified by key works such as the *Principal Odour Map* [89] and *eigengraphs* [135].

Olfaction in Static Scenes. This paradigm addresses tasks where AI systems leverage olfaction in conjunction with other modalities (primarily vision) to understand and interact with relatively stable scenes. It focuses on: (a) Detection tasks (akin to visual grounding tasks [56, 92]) such as *olfactory-visual localization* in static scenes where we focus on pinpointing the specific object or area within a static 2D/3D scene that is the source of a detected odour, using both visual cues and olfactory sensor data. For example, identifying which container in a pantry is emitting a specific smell or which fruit in a bowl is ripe. (b) Reasoning tasks (akin to traditional video question answering [151, 73]) such as *olfactory-visual question answering* where we focus on answering questions about a static scenes that require joint reasoning over visual and olfactory information. For example, "What ingredients are likely present in this (unseen) dish based on its smell and the visible cooking setup?"

Olfaction in Dynamic Scenes. This paradigm focuses on tasks involving mobile agents, changing olfactory landscapes, and the temporal evolution of scents, requiring robust integration of olfactory data with spatio-temporal visual and potentially auditory information. It focuses on: (a) Localization & Tracking tasks (akin to multi-object tracking [42]) where we focus on detecting, localizing, and tracking the source(s) of odours as they move. This could involve tracking a scent plume to its origin or identifying a moving entity by its olfactory signature. (b) Navigation tasks (akin to visual navigation [78]) where we enable an agent to navigate complex, potentially visually ambiguous environments by primarily relying on or being significantly aided by olfactory cues. This includes tasks like navigating to a specific odour source, using a sequence of olfactory landmarks, or avoiding areas with hazardous smells [24, 25, 54, 61]. Current robots often struggle to effectively use chemical cues [25]. New benchmarks are needed to test an AI's proficiency in pinpointing odour sources, such as a gas leak in a building [24, 54], tracking dynamic scent plumes [131], or navigating using sparse olfactory landmarks [40] (c) Reasoning tasks (akin to video question answering [67, 153], multimodal event understanding [57, 110]) where focus on answering complex questions or generating summaries about dynamic events and activities by integrating information from olfactory, visual (video), and auditory streams. For example, "Based on the sight of smoke, the sound of a crackling fire, and the smell of burning wood, what is likely happening and where?"

The creation of these benchmarks is not an academic exercise; it is a **critical enabler** for pushing Embodied AI towards genuine environmental understanding and interaction. Addressing the current void, exemplified by the absence of olfactory components in general AI assessments like the ARC-AGI benchmark [31], is essential. Only by rigorously testing against such benchmarks can we ensure that machine olfaction evolves from a niche curiosity into a core competency of embodied systems.

3 Ethical Considerations in Superhuman Olfaction

History has shown that AI systems which achieve superhuman performance in narrow modalities can yield unintended and sometimes adverse consequences. In computer vision, superhuman accuracy in facial recognition has led to widespread concerns about surveillance, privacy, and algorithmic bias, particularly when deployed without consent or adequate regulation [23, 118]. In NLP, large language models have demonstrated uncanny fluency but have also amplified disinformation, toxicity, and epistemic bias at scale [15, 146]. In audition, AI-generated voices now impersonate individuals with near-perfect fidelity—blurring the boundary between legitimate media and deepfakes [102].

As we now move toward endowing machines with olfactory intelligence—particularly one that exceeds human capabilities—we must anticipate similar dual-use risks. Superhuman olfaction could be used to enhance environmental monitoring, disease diagnostics, or search-and-rescue robotics. But it could equally be leveraged for mass surveillance of human biological states (e.g., stress, fertility, intoxication), infringing on bodily privacy in ways for which no legal or ethical precedent currently exists [70]. Just as one's voice and biometrics are considered personal identifiable information (PII), so is it also probable that one's olfactory characteristics through their breath [81, 12, 11] and general aroma data may also be classified as PII in the near future.

The ethical stakes escalate when these narrow superhuman modalities are integrated into unified multimodal architectures. Emerging olfaction-vision-language models (OVLMs) [60, 127] combine chemical sensing with high-fidelity perception and reasoning systems. Under the wrong intentions, these models could infer behavioral, physiological, and even psychological states of individuals with unprecedented granularity. Biased olfactory agents could guide humans to harmful chemicals, or let spoiled or poisoned ingredients pass quality control within a cosmetics manufacturing line. Merging olfactory intelligence with large multimodal foundation models grants one more super-human sense to machines that already out-perform humans on many tasks [30]. Embodying olfaction in robotic form gives AI a means of exploring the world with an additional medium with which to further understand multimodal patterns. Yet it must be understood that the additional inference capabilities could be exploited for commercial manipulation, targeted policing, or covert tracking of marginalized populations. The convergence of modalities does not dilute ethical risks; it compounds them.

AI ethics frameworks to date have been disproportionately focused on vision, language, and fairness metrics rooted in social identity categories defined by vision, audition, and language (e.g., gender, race, dialect). These frameworks are not yet equipped to handle new sensory axes such as scent, which interact with human dignity, consent, and neurophysiological privacy in complex ways. There is currently no consensus on what it means to "explain" a decision made by an olfactory model or to audit bias in an odour classifier trained on population-specific olfaction data.

4 Discussion, Limitations & Future Work

The lack of large datasets in olfaction precludes the use of many modern "big data" machine learning techniques. Adaptive [156], continuous [63], and few-shot learning [144] techniques will most certainly need to be employed in the near term to progress olfactory intelligence. Co-training [18, 95, 94, 62] and federated learning [100] methods can help produce more confident datasets over smaller disjoint datasets. Evidential methods based on the Dempster-Shafer Theory [128, 4] can help seed much larger datasets from empirical evidence gathered from smaller confident samples. The temporal and information sparsity of olfaction data [48] (especially during navigation) may indeed lead to adaptive architectures that allow the model to be modified as it trains through methods such as columnar constructive networks [77] and prototypical networks [132].

Solving these key challenges within olfaction enables more applications such as scent-based navigation, more accurate medical diagnostics, organic improvements to agriculture, and better quality control of consumer products (see Section B of the Supplementary Material for more detail on potential applications). The generation of new aromas and even commercial products [145] centered around novel odours become possible by combining generative AI techniques with olfaction [59, 106].

As the state of olfaction progresses, we expect to see many changes around hardware, software, and overall thought leadership. Sensors will continue to increase in detection speed [46, 74, 103] and resolution [134] and decrease in size. Many of these optimizations will come through improvements of materials science and sensor manufacturing of which several opportunities exist [13, 39]. Electronic noses will become more integrated into society as the potential applications become realized and olfactory sensors approach the ubiquity of cameras and microphones. The proliferation of such sensors will allow us to materialize the dense data available in the air around us. Momentum in artificial olfaction will motivate progress in neuroscience to allow us to better understand aspects of the brain still unsolved [5, 14, 101, 113, 129, 155].

Our work above highlights significant opportunities to make a monumental step in AI. Progressing artificial olfaction will bring the state of AI closer to truly embodied intelligence by allowing robots to navigate and perceive the world with an additional high-context dimension. The fields of AI and robotics are lacking one key sense that is significantly hindering their advancement. We hope our work here raises awareness about the plethora of opportunities that exist in the field of olfaction and motivates funding, development, and research in the field on par with those received by other modalities. Convergence on the five key problems presented here will bring us one step closer to enabling the sense of smell for machines.

References

[1] Tobias Ackels, Andrew Erskine, Debanjan Dasgupta, Alina Cristina Marin, Tom P.A. Warner, Sina Tootoonian, Izumi Fukunaga, Julia J. Harris, and Andreas T. Schaefer. Fast odour dynamics are encoded in the olfactory system and guide behaviour. *Nature*, 593(7860):558–563, 2021.

[2] Ibrahim M Alabdulmohsin, Behnam Neyshabur, and Xiaohua Zhai. Revisiting neural scaling laws in language and vision. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 22300–22312. Curran Associates, Inc., 2022.

[3] Ebtsam K. Alenezy, Ahmad E. Kandjani, Mahdokht Shaibani, Adrian Trinchi, Suresh K. Bhargava, Samuel J. Ippolito, and Ylias Sabri. Human breath analysis; clinical application and measurement: An overview. *Biosensors and Bioelectronics*, 278:117094, 2025.

[4] Alexander Amini, Wilko Schwarting, Ava Soleimany, and Daniela Rus. Deep evidential regression. *Advances in Neural Information Processing Systems*, 33, 2020.

[5] Erica F. Andrews, Raluca Pascalau, Alexandra Horowitz, Gillian M. Lawrence, and Philippa J. Johnson. Extensive connections of the canine olfactory pathway revealed by tractography and dissection. *Journal of Neuroscience*, 42(33):6392–6407, 2022.

[6] Aryballe. Automotive cabin odors, n.d. Accessed: 2025-05-23.

[7] Aryballe. Digital olfaction & fermentation: What you need to know, n.d. Accessed: 2025-05-23.

[8] Ariel Flint Ashery, Luca Maria Aiello, and Andrea Baronchelli. Emergent social conventions and collective bias in llm populations. *Science Advances*, 11(20):eadu9368, 2025.

[9] Leffingwell & Associates. Pmp 2001 - database of perfumery materials and performance. `http://www.leffingwell.com/bacispmp.htm`. Accessed: 2025-03-08.

[10] Jibo Bai, Baojiang Li, Xichao Wang, Haiyan Wang, and Yuting Guo. Bionic hand motion control method based on imitation of human hand movements and reinforcement learning. *Journal of Bionic Engineering*, 21(2):764–777, Mar 2024.

[11] Ivneet Banga, Kordel France, Anirban Paul, and Shalini Prasad. E.co.tech breathalyzer: A pilot study of a non-invasive covid-19 diagnostic tool for light and non-smokers. *ACS Measurement Science Au*, 4(5):496–503, Oct 2024.

[12] Ivneet Banga, Anirban Paul, Kordel France, Ben Micklich, Bret Cardwell, Craig Micklich, and Shalini Prasad. E.co.tech-electrochemical handheld breathalyzer covid sensing technology. *Scientific Reports*, 12(1):4370, 2022.

[13] Ahmed Barhoum, Selma Hamimed, Hamda Slimi, Amina Othmani, Fatehy M. Abdel-Haleem, and Mikhael Bechelany. Modern designs of electrochemical sensor platforms for environmental analyses: Principles, nanofabrication opportunities, and challenges. *Trends in Environmental Analytical Chemistry*, 38:e00199, 2023.

[14] Ann-Sophie Barwich. Olfaction is a spatial sense. *Review of Philosophy and Psychology*, Jan 2025.

[15] Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 610–623, 2021.

[16] Christian B. Billesbølle, Claire A. de March, Wijnand J. C. van der Velden, Ning Ma, Jeevan Tewari, Claudia Llinas del Torrent, Linus Li, Bryan Faust, Nagarajan Vaidehi, Hiroaki Matsunami, and Aashish Manglik. Structural basis of odorant recognition by a human odorant receptor. *Nature*, 615(7953):742–749, Mar 2023.

[17] Eric Block, Seogjoo Jang, Hiroaki Matsunami, Sivakumar Sekharan, Bérénice Dethier, Mehmed Z Ertem, Sivaji Gundala, Yi Pan, Shengju Li, Zhen Li, Stephene N Lodge, Mehmet Ozbil, Huihong Jiang, Sonia F Penalba, Victor S Batista, and Hanyi Zhuang. Implausibility of the vibrational theory of olfaction. *Proceedings of the National Academy of Sciences of the United States of America*, 112(21):E2766–E2774, May 2015. Research Support, N.I.H., Extramural; Research Support, Non-U.S. Gov't; Research Support, U.S. Gov't, Non-P.H.S.

[18] Avrim Blum and Tom Mitchell. Combining labeled and unlabeled data with co-training. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, COLT' 98, pages 92–100, New York, NY, USA, 1998. Association for Computing Machinery.

[19] Boston Dynamics. Atlas, n.d. Accessed: 2025-05-23.

[20] A. E. Bourgeois and Joanne O. Bourgeois. Theories of olfaction: A review. *Revista Interamericana de Psicología/Interamerican Journal of Psychology*, 4(1), Jul. 2017.

[21] Gregory N. Bratman, Cecilia Bembibre, Gretchen C. Daily, Richard L. Doty, Thomas Hummel, Lucia F. Jacobs, Peter H. Kahn, Connor Lashus, Asifa Majid, John D. Miller, Anna Oleszkiewicz, Hector Olvera-Alvarez, Valentina Parma, Anne M. Riederer, Nancy Long Sieber, Jonathan Williams, Jieling Xiao, Chia-Pin Yu, and John D. Spengler. Nature and human well-being: The olfactory pathway. *Science Advances*, 10(20):eadn3028, 2024.

[22] Jennifer C Brookes, Filio Hartoutsiou, AP Horsfield, and AM Stoneham. Could humans recognize odor by phonon assisted tunneling? *Physical review letters*, 98(3):038101, 2007.

[23] Joy Buolamwini and Timnit Gebru. Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of the Conference on Fairness, Accountability and Transparency (FAT)*, pages 77–91, 2018.

[24] Javier Burgués, Victor Hernández, Achim J. Lilienthal, and Santiago Marco. Smelling nano aerial vehicle for gas source localization and mapping. *Sensors*, 19(3), 2019.

[25] Javier Burgués and Santiago Marco. Environmental chemical sensing using small drones: A review. *Science of The Total Environment*, 748:141172, 2020.

[26] Shawn D Burton, Audrey Brown, Thomas P Eiting, Isaac A Youngstrom, Thomas C Rust, Michael Schmuker, and Matt Wachowiak. Mapping odorant sensitivities reveals a sparse but structured representation of olfactory chemical space by sensory input to the mouse olfactory bulb. *eLife*, 11:e80470, July 2022.

[27] Antonio Celani, Emmanuel Villermaux, and Massimo Vergassola. Odor Landscapes in Turbulent Environments. *Physical Review X*, 4(4):041015, October 2014.

[28] Minqi Chen, Zhongqian Song, Shengjie Liu, Zhenbang Liu, Weiyan Li, Huijun Kong, Cong Li, Yu Bao, Wei Zhang, and Li Niu. Iontronic tactile sensory system for plant species and growth-stage classification. *Device*, 3(3), Mar 2025.

[29] Cheng Chi, Siyuan Feng, Yilun Du, Zhenjia Xu, Eric Cousineau, Benjamin Burchfiel, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. In *Proceedings of Robotics: Science and Systems (RSS)*, 2023.

[30] Wei-Lin Chiang, Lianmin Zheng, Ying Sheng, Anastasios Nikolas Angelopoulos, Tianle Li, Dacheng Li, Hao Zhang, Banghua Zhu, Michael Jordan, Joseph E. Gonzalez, and Ion Stoica. Chatbot arena: An open platform for evaluating llms by human preference, 2024.

[31] François Chollet. On the measure of intelligence, 2019.

[32] M. A. Z. Chowdhury and M. A. Oehlschlaeger. Artificial intelligence in gas sensing: A review. *ACS Sensors*, 10(3):1538–1563, Mar 2025.

[33] Clone Robotics. Clone robotics, n.d. Accessed: 2025-05-23.

[34] Open X-Embodiment Collaboration, Abby O'Neill, Abdul Rehman, Abhinav Gupta, and et al. Open x-embodiment: Robotic learning datasets and rt-x models, 2025.

[35] The Good Scents Company. The good scents company information system. http://www.thegoodscentscompany.com/. Accessed: 2025-03-08.

[36] Astrid R Concha, Claire M Guest, Rob Harris, Thomas W Pike, Alexandre Feugier, Helen Zulch, and Daniel S Mills. Canine olfactory thresholds to amyl acetate in a biomedical detection scenario. *Frontiers in Veterinary Science*, 5:345, 2019.

[37] Mark Conway, Merve Oncul, Kate Allen, Zongqian Zhang, and Jamie Johnston. Perceptual constancy for an odor is acquired through changes in primary sensory neurons. *Science Advances*, 10(50):eado9205, December 2024.

[38] Cornell University. arxiv api. https://info.arxiv.org/help/api/index.html, 2025. Accessed: 2025-05-16.

[39] James A. Covington, Santiago Marco, Krishna C. Persaud, Susan S. Schiffman, and H. Troy Nagle. Artificial olfaction in the 21st century. *IEEE Sensors Journal*, 21(11):12969–12990, June 2021.

[40] John Crimaldi, Hong Lei, Andreas Schaefer, Michael Schmuker, Brian H. Smith, Aaron C. True, Justus V. Verhagen, and Jonathan D. Victor. Active sensing in a dynamic olfactory world. *Journal of Computational Neuroscience*, 50(1):1–6, Feb 2022.

[41] Angela Dai, Matthias Nießner, Michael Zollöfer, Shahram Izadi, and Christian Theobalt. Bundlefusion: Real-time globally consistent 3d reconstruction using on-the-fly surface re-integration. *ACM Transactions on Graphics 2017 (TOG)*, 2017.

[42] Achal Dave, Tarasha Khurana, Pavel Tokmakov, Cordelia Schmid, and Deva Ramanan. Tao: A large-scale benchmark for tracking any object. In *European Conference on Computer Vision*, 2020.

[43] DeepSeek-AI. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning, 2025.

[44] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.

[45] Zelin Deng, Yunlong Dong, and Xing Liu. Reward shaping in reinforcement learning for robotic hand manipulation. *Neurocomputing*, 638:130204, 2025.

[46] Nik Dennler, Damien Drix, Tom PA Warner, Shavika Rastogi, Cecilia Della Casa, Tobias Ackels, Andreas T Schaefer, André van Schaik, and Michael Schmuker. High-speed odor sensing using miniaturized electronic nose. *Science Advances*, 10(45):eadp1764, 2024.

[47] Nik Dennler, Shavika Rastogi, Jordi Fonollosa, André van Schaik, and Michael Schmuker. Drift in a popular metal oxide sensor dataset reveals limitations for gas classification benchmarks. *Sensors and Actuators B: Chemical*, 361:131668, 2022.

[48] Nik Dennler, Aaron True, André van Schaik, and Michael Schmuker. Neuromorphic principles for machine olfaction. *Neuromorphic Computing and Engineering*, 5(2):023001, may 2025.

[49] Nik Dennler, André van Schaik, and Michael Schmuker. Limitations in odour recognition and generalization in a neuromorphic olfactory circuit. *Nature Machine Intelligence*, 6:1451–1453, 2024.

[50] Vikram Narayanan Dhamu, Anirban Paul, Sriram Muthukumar, and Shalini Prasad. Electrochemical framework for dynamic tracking of soil organic matter. *Biosensors and Bioelectronics: X*, 17:100440, 2024.

[51] Marta Díaz-Quesada, Isaac A Youngstrom, Yusuke Tsuno, Kyle R Hansen, Michael N Economo, and Matt Wachowiak. Inhalation frequency controls reformatting of mitral/tufted cell odor representations in the olfactory bulb. *Journal of Neuroscience*, 38(9):2189–2206, 2018.

[52] Richard L Doty and Vidyul Kamath. Smell identification ability: changes with age. *Science of Aging Knowledge Environment*, 2001(10):cp2, 2001.

[53] Richard L Doty, Peter Shaman, Steven L Applebaum, Roland Giberson, Lisa Siksorski, and L Rosenberg. Influence of age and age-related diseases on olfactory function. *Annals of the New York Academy of Sciences*, 435(1):36–49, 1984.

[54] Bardienus P. Duisterhof, Shushuai Li, Javier Burgués, Vijay Janapa Reddi, and Guido C. H. E. de Croon. Sniffy bug: A fully autonomous swarm of gas-seeking nano quadcopters in cluttered environments. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 9099–9106, 2021.

[55] GM Dyson. Some aspects of the vibration theory of odor. *Perfumery and essential oil record*, 19(456-459), 1928.

[56] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, 88(2):303–338, June 2010.

[57] Fadime Sener, Dibyadip Chatterjee, Daniel Shelepov, Kun He, Dipika Singhania, Robert Wang, and Angela Yao. Assembly101: A Large-Scale Multi-View Video Dataset for Understanding Procedural Activities. *Computer Vision and Pattern Recognition*, 2022.

[58] Figure AI. Figure ai, n.d. Accessed: 2025-05-23.

[59] James Fong, Hannah K. Doyle, Congli Wang, Alexandra E. Boehm, Sofie R. Herbeck, Vimal Prabhu Pandiyan, Brian P. Schmidt, Pavan Tiruveedhula, John E. Vanston, William S. Tuten, Ramkumar Sabesan, Austin Roorda, and Ren Ng. Novel color via stimulation of individual photoreceptors at population scale. *Science Advances*, 11(16):eadu1052, 2025.

[60] Kordel K. France. Scentience. `https://apps.apple.com/us/app/scentience/id6741092923`, 2025. Accessed: 2025-03-02.

[61] Kordel K. France, Anirban Paul, Ivneet Banga, and Shalini. Prasad. Emergent behavior in evolutionary swarms for machine olfaction. In *Proceedings of the Genetic and Evolutionary Computation Conference*, GECCO '24, page 30–38, New York, NY, USA, 2024. Association for Computing Machinery.

[62] Kordel K. France and John W. Sheppard. Factored particle swarm optimization for policy co-training in reinforcement learning. In *Proceedings of the Genetic and Evolutionary Computation Conference*, GECCO '23, page 30–38, New York, NY, USA, 2023. Association for Computing Machinery.

[63] Yunhao Ge, Yuecheng Li, Di Wu, Ao Xu, Adam M. Jones, Amanda Sofie Rios, Iordanis Fostiropoulos, Shixian Wen, Po-Hsuan Huang, Zachary William Murdock, Gozde Sahin, Shuo Ni, Kiran Lekkala, Sumedh Anand Sontakke, and Laurent Itti. Lightweight learner for shared knowledge lifelong learning, 2023.

[64] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.

[65] Ruben Grandia, Espen Knoop, Michael Hopkins, Georg Wiedebach, Jared Bishop, Steven Pickles, David Müller, and Moritz Bächer. Design and control of a bipedal robotic character. In *Robotics: Science and Systems XX*, RSS2024. Robotics: Science and Systems Foundation, July 2024.

[66] Kristen Grauman, Andrew Westbury, Eugene Byrne, Zachary Chavis, Antonino Furnari, Rohit Girdhar, Jackson Hamburger, Hao Jiang, Miao Liu, Xingyu Liu, Miguel Martin, Tushar Nagarajan, Ilija Radosavovic, Santhosh Kumar Ramakrishnan, Fiona Ryan, Jayant Sharma, Michael Wray, Mengmeng Xu, Eric Zhongcong Xu, Chen Zhao, Siddhant Bansal, Dhruv Batra, Vincent Cartillier, Sean Crane, Tien Do, Morrie Doulaty, Akshay Erapalli, Christoph Feichtenhofer, Adriano Fragomeni, Qichen Fu, Abrham Gebreselasie, Cristina Gonzalez, James Hillis, Xuhua Huang, Yifei Huang, Wenqi Jia, Weslie Khoo, Jachym Kolar, Satwik

Kottur, Anurag Kumar, Federico Landini, Chao Li, Yanghao Li, Zhenqiang Li, Karttikeya Mangalam, Raghava Modhugu, Jonathan Munro, Tullie Murrell, Takumi Nishiyasu, Will Price, Paola Ruiz Puentes, Merey Ramazanova, Leda Sari, Kiran Somasundaram, Audrey Southerland, Yusuke Sugano, Ruijie Tao, Minh Vo, Yuchen Wang, Xindi Wu, Takuma Yagi, Ziwei Zhao, Yunyi Zhu, Pablo Arbelaez, David Crandall, Dima Damen, Giovanni Maria Farinella, Christian Fuegen, Bernard Ghanem, Vamsi Krishna Ithapu, C. V. Jawahar, Hanbyul Joo, Kris Kitani, Haizhou Li, Richard Newcombe, Aude Oliva, Hyun Soo Park, James M. Rehg, Yoichi Sato, Jianbo Shi, Mike Zheng Shou, Antonio Torralba, Lorenzo Torresani, Mingfei Yan, and Jitendra Malik. Ego4D: Around the World in 3,000 Hours of Egocentric Video. *arXiv*, October 2021.

[67] Madeleine Grunde-McLaughlin, Ranjay Krishna, and Maneesh Agrawala. Agqa 2.0: An updated benchmark for compositional spatio-temporal reasoning, 2022.

[68] R. Gutierrez-Osuna. Pattern analysis for machine olfaction: a review. *IEEE Sensors Journal*, 2(3):189–202, 2002.

[69] Tuomas Haarnoja, Sehoon Ha, Aurick Zhou, Jie Tan, George Tucker, and Sergey Levine. Learning to walk via deep reinforcement learning. In *Proceedings of Robotics: Science and Systems*, FreiburgimBreisgau, Germany, June 2019.

[70] Rosemary Hains, Noah Overman, and Ann-Sophie Barwich. The ethics of olfactory surveillance: A nascent area of inquiry. *AI & Society*, 36:841–850, 2021.

[71] Abdallah Hayar, Sergei Karnup, Michael T Shipley, and Matthew Ennis. Olfactory bulb glomeruli: external tufted cells intrinsically burst at theta frequency and are entrained by patterned olfactory input. *Journal of Neuroscience*, 24(5):1190–1199, 2004.

[72] Tajana Horvat, Gordana Pehnec, and Ivana Jakovljević. Volatile organic compounds in indoor air: Sampling, determination, sources, health risk, and regulatory insights. *Toxics*, 13(5), 2025.

[73] Drew A Hudson and Christopher D Manning. Gqa: A new dataset for real-world visual reasoning and compositional question answering. *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[74] Nabil Imam and Thomas A. Cleland. Rapid online learning and robust recall in a neuromorphic olfactory circuit. *Nature Machine Intelligence*, 2(3):181–191, March 2020.

[75] Peter. J Irga, Gabrielle Mullen, Robert Fleck, Stephen Matheson, Sara. J Wilkinson, and Fraser. R Torpy. Volatile organic compounds emitted by humans indoors– a review on the measurement, test conditions, and analysis techniques. *Building and Environment*, 255:111442, 2024.

[76] Md Ahmadul Islam and Shahidul Islam. Sourdough bread quality: Facts and factors. *Foods*, 13(13), July 2024.

[77] Khurram Javed, Haseeb Shah, Richard S. Sutton, and Martha White. Scalable real-time recurrent learning using columnar-constructive networks. *Journal of Machine Learning Research*, 24(256):1–34, 2023.

[78] Faith Johnson, Bryan Bo Cao, Kristin Dana, Shubham Jain, and Ashwin Ashok. A landmark-aware visual navigation dataset, 2025.

[79] Tobias Jost, Thomas Heymann, and Marcus A Glomb. Efficient analysis of 2-acetyl-1-pyrroline in foods using a novel derivatization strategy and LC-MS/MS. *J Agric Food Chem*, 67(10):3046–3054, February 2019.

[80] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, Aug 2021.

[81] Maria Kaloumenou, Evangelos Skotadis, Nefeli Lagopati, Efstathios Efstathopoulos, and Dimitris Tsoukalas. Breath analysis: A promising tool for disease diagnosis—the role of sensors. *Sensors*, 22(3), 2022.

[82] Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B. Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu, and Dario Amodei. Scaling laws for neural language models, 2020.

[83] Gunay Kazimzade and Milagros Miceli. Biased priorities, biased outcomes: Three recommendations for ethics-oriented data annotation practices. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, AIES '20, page 71, New York, NY, USA, 2020. Association for Computing Machinery.

[84] Andreas Keller and Leslie B. Vosshall. Olfactory perception of chemically diverse molecules. *BMC Neuroscience*, 17(1):55, Aug 2016.

[85] Chuntae Kim, Kyung Kwan Lee, Moon Sung Kang, Dong-Myeong Shin, Jin-Woo Oh, Chang-Soo Lee, and Dong-Wook Han. Artificial olfactory sensor technology that mimics the olfactory mechanism: a comprehensive review. *Biomaterials Research*, 26(1):40, 2022.

[86] Rodney Michael Kinney, Chloe Anastasiades, Russell Authur, Iz Beltagy, Jonathan Bragg, Alexandra Buraczynski, Isabel Cachola, Stefan Candra, Yoganand Chandrasekhar, Arman Cohan, Miles Crawford, Doug Downey, Jason Dunkelberger, Oren Etzioni, Rob Evans, Sergey Feldman, Joseph Gorney, David W. Graham, F.Q. Hu, Regan Huff, Daniel King, Sebastian Kohlmeier, Bailey Kuehl, Michael Langan, Daniel Lin, Haokun Liu, Kyle Lo, Jaron Lochner, Kelsey MacMillan, Tyler C. Murray, Christopher Newell, Smita R Rao, Shaurya Rohatgi, Paul Sayre, Zejiang Shen, Amanpreet Singh, Luca Soldaini, Shivashankar Subramanian, A. Tanaka, Alex D Wade, Linda M. Wagner, Lucy Lu Wang, Christopher Wilhelm, Caroline Wu, Jiangjiang Yang, Angele Zamarron, Madeleine van Zuylen, and Daniel S. Weld. The semantic scholar open data platform. *ArXiv*, abs/2301.10140, 2023.

[87] Jan-Christoph Klie, Richard Eckart de Castilho, and Iryna Gurevych. Analyzing dataset annotation quality management in the wild, 2024.

[88] Dhireesha Kudithipudi, Catherine Schuman, Craig M. Vineyard, Tej Pandit, Cory Merkel, Rajkumar Kubendran, James B. Aimone, Garrick Orchard, Christian Mayr, Ryad Benosman, Joe Hays, Cliff Young, Chiara Bartolozzi, Amitava Majumdar, Suma George Cardwell, Melika Payvand, Sonia Buckley, Shruti Kulkarni, Hector A. Gonzalez, Gert Cauwenberghs, Chetan Singh Thakur, Anand Subramoney, and Steve Furber. Neuromorphic computing at scale. *Nature*, 637(8047):801–812, January 2025.

[89] Brian K. Lee, Emily J. Mayhew, Benjamin Sanchez-Lengeling, Jennifer N. Wei, Wesley W. Qian, Kelsie A. Little, Matthew Andres, Britney B. Nguyen, Theresa Moloy, Jacob Yasonik, Jane K. Parker, Richard C. Gerkin, Joel D. Mainland, and Alexander B. Wiltschko. A principal odor map unifies diverse tasks in olfactory perception. *Science*, 381(6661):999–1006, 2023.

[90] Gregory Leonardos, David Kendall, and Nancy Barnard. Odor threshold determination of 53 odorant chemicals. *Journal of Environmental Conservation Engineering*, 3(8):579–585, 1974.

[91] Cheng Li, Kun Nie, Yan Ma, Mingming Shao, and Guangjun Li. Olfactory dysfunction in parkinson's disease: a systematic review and meta-analysis. *Frontiers in Aging Neuroscience*, 8:145, 2016.

[92] Tsung-Yi Lin, Michael Maire, Serge J. Belongie, Lubomir D. Bourdev, Ross B. Girshick, James Hays, Pietro Perona, Deva Ramanan, Piotr Doll'a r, and C. Lawrence Zitnick. Microsoft COCO: common objects in context. *CoRR*, abs/1405.0312, 2014.

[93] Friedrich Longin, Heiner Beck, Hermann Gütler, Wendelin Heilig, Michael Kleinert, Matthias Rapp, Norman Philipp, Alexander Erban, Dominik Brilhaus, Tabea Mettler-Altmann, and Benjamin Stich. Aroma and quality of breads baked from old and modern wheat varieties and their prediction from genomic and flour-based metabolite profiles. *Food Research International*, 129:108748, 2020.

[94] Fan Ma, Deyu Meng, Xuanyi Dong, and Yi Yang. Self-paced multi-view co-training. *Journal of Machine Learning Research*, 21(57):1–38, 2020.

[95] Fan Ma, Deyu Meng, Qi Xie, Zina Li, and Xuanyi Dong. Self-paced co-training. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 2275–2284. PMLR, 06–11 Aug 2017.

[96] Limei Ma, Qiang Qiu, Stephen Gradwohl, Aaron Scott, Elden Q. Yu, Richard Alexander, Winfried Wiegraebe, and C. Ron Yu. Distributed representation of chemical features and tunotopic organization of glomeruli in the mouse olfactory bulb. *Proceedings of the National Academy of Sciences*, 109(14):5481–5486, April 2012.

[97] Joel D Mainland, Andreas Keller, Ying Li, Ting Zhou, Christian Trimmer, Lawrence L Snyder, Angela H Moberly, Kayla A Adipietro, Wei L Liu, Hong Zhuang, et al. The missense of smell: functional variability in the human odorant receptor repertoire. *Nature Neuroscience*, 17(1):114–120, 2014.

[98] G. Malcolm Dyson. The scientific basis of odour. *Journal of the Society of Chemical Industry*, 57(28):647–651, 1938.

[99] Bettina Malnic, Junzo Hirono, Takaaki Sato, and Linda B Buck. Combinatorial receptor codes for odors. *Cell*, 96(5):713–723, 1999.

[100] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-Efficient Learning of Deep Networks from Decentralized Data. In Aarti Singh and Jerry Zhu, editors, *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pages 1273–1282. PMLR, 20–22 Apr 2017.

[101] Eduardo Mercado, 3rd and Jessica Zhuo. Do rodents smell with sound? *Neurosci Biobehav Rev*, 167:105908, September 2024.

[102] Thorsten Meyer and Mads Christensen. Deepfake voice synthesis: A threat to privacy, security, and trust in voice communication. *IEEE Signal Processing Magazine*, 39(1):99–106, 2022.

[103] Md. Abdul Momin, Masaya Toda, Zhuqing Wang, Mai Yamazaki, Krzysztof Moorthi, Yasuaki Kawaguchi, and Takahito Ono. Investigation towards nanomechanical sensor array for real-time detection of complex gases. *Microsystems & Nanoengineering*, 11(1):53, Mar 2025.

[104] Daniel Münch and C Giovanni Galizia. Door 2.0-comprehensive mapping of drosophila melanogaster odorant responses. *Scientific reports*, 6(1):21841, 2016.

[105] Daniel Münch and C. Giovanni Galizia. DoOR 2.0 - Comprehensive Mapping of Drosophila melanogaster Odorant Responses. *Scientific Reports*, 6(1):21841, February 2016.

[106] K. Nakamoto. *Digital Technologies in Olfaction*. Elsevier, 2023.

[107] Yoshihito Niimura and Masatoshi Nei. Evolutionary dynamics of olfactory and other chemosensory receptor genes in vertebrates. *Journal of human genetics*, 51(6):505–517, 2006.

[108] Osmo AI. Generation by osmo, n.d. Accessed: 2025-05-23.

[109] Osmo AI. Osmo ai, n.d. Accessed: 2025-05-23.

[110] Rohith Peddi, Shivvrat Arya, Bharath Challa, Likhitha Pallapothula, Akshay Vyas, Bhavya Gouripeddi, Qifan Zhang, Jikai Wang, Vasundhara Komaragiri, Eric Ragan, Nicholas Ruozzi, Yu Xiang, and Vibhav Gogate. Captaincook4d: A dataset for understanding errors in procedural activities. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, editors, *Advances in Neural Information Processing Systems*, volume 37, pages 135626–135679. Curran Associates, Inc., 2024.

[111] Karol J. Piczak. ESC: Dataset for Environmental Sound Classification. In *Proceedings of the 23rd Annual ACM Conference on Multimedia*, pages 1015–1018. ACM Press.

[112] Tianfei Qi, Ruonan Jia, Linbo Zhang, Guanyu Li, Xuefeng Wei, Jianting Ye, Ruilian Qi, and Haiyuan Chen. Real-time and visual monitoring coating damage and metal corrosion based on the multiple chromatic transition of tmb triggered by highly active n-cds/fe3o4 peroxidase mimics and fe3+. *Sensors and Actuators B: Chemical*, 415:136003, 2024.

[113] Wesley W Qian, Jennifer N Wei, Benjamin Sanchez-Lengeling, Brian K Lee, Yunan Luo, Marnix Vlot, Koen Dechering, Jian Peng, Richard C Gerkin, and Alexander B Wiltschko. Metabolic activity organizes olfactory representations. *eLife*, 12:e82502, may 2023.

[114] Alec Radford, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. Robust speech recognition via large-scale weak supervision, 2022.

[115] Ilija Radosavovic, Tete Xiao, Bike Zhang, Trevor Darrell, Jitendra Malik, and Koushil Sreenath. Real-world humanoid locomotion with reinforcement learning. *Science Robotics*, 9(89):eadi9579, 2024.

[116] Bipin Rajendran, Abu Sebastian, Michael Schmuker, Narayan Srinivasa, and Evangelos Eleftheriou. Low-power neuromorphic hardware for signal processing applications: A review of architectural and system-level design approaches. *IEEE Signal Processing Magazine*, 36(6):97–110, 2019.

[117] Bipin Rajendran, Abu Sebastian, Michael Schmuker, Narayan Srinivasa, and Evangelos Eleftheriou. Low-power neuromorphic hardware for signal processing applications: A review of architectural and system-level design approaches. *IEEE Signal Processing Magazine*, 36(6):97–110, 2019.

[118] Inioluwa Deborah Raji, Joy Buolamwini, Solon Barocas, Matt Krishnan, and Julian Lemos. Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, pages 429–435, 2020.

[119] Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. Squad: 100,000+ questions for machine comprehension of text, 2016.

[120] Shavika Rastogi, Nik Dennler, Michael Schmuker, and André van Schaik. The neuromorphic analog electronic nose. *ArXiv*, abs/2410.16677, 2024.

[121] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models, 2022.

[122] Alex D Rygg, Blaire Van Valkenburgh, and Brent A Craven. The influence of sniffing on airflow and odorant deposition in the canine nasal cavity. *Chemical senses*, 42(8):683–698, 2017.

[123] Khaled Saab, Jan Freyberg, Chunjong Park, Tim Strother, Yong Cheng, Wei-Hung Weng, David G. T. Barrett, David Stutz, Nenad Tomasev, Anil Palepu, Valentin Liévin, Yash Sharma, Roma Ruparel, Abdullah Ahmed, Elahe Vedadi, Kimberly Kanada, Cian Hughes, Yun Liu, Geoff Brown, Yang Gao, Sean Li, S. Sara Mahdavi, James Manyika, Katherine Chou, Yossi Matias, Avinatan Hassidim, Dale R. Webster, Pushmeet Kohli, S. M. Ali Eslami, Joëlle Barral, Adam Rodman, Vivek Natarajan, Mike Schaekermann, Tao Tu, Alan Karthikesalingam, and Ryutaro Tanno. Advancing conversational diagnostic ai with multimodal reasoning, 2025.

[124] J. Salamon, C. Jacoby, and J. P. Bello. A dataset and taxonomy for urban sound research. In *22nd ACM International Conference on Multimedia (ACM-MM'14)*, pages 1041–1044, Orlando, FL, USA, Nov. 2014.

[125] Nithya Sambasivan, Shivani Kapania, Hannah Highfill, Diana Akrong, Praveen Paritosh, and Lora M Aroyo. "everyone wants to do the model work, not the data work": Data cascades in high-stakes ai. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.

[126] Benjamin Sanchez-Lengeling, Jennifer N. Wei, Brian K. Lee, Richard C. Gerkin, Alán Aspuru-Guzik, and Alexander B. Wiltschko. Machine learning for scent: Learning generalizable perceptual representations of small molecules, 2019.

[127] Scentience. Scentience app: World's first olfaction-vision language model. 2025. Accessed: 2025-03-02.

[128] Murat Sensoy, Lance Kaplan, and Melih Kandemir. Evidential deep learning to quantify classification uncertainty. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

[129] Aditi Seshadri, Heta A Gandhi, Geemi P Wellawatte, and Andrew D White. Why does that molecule smell? December 2022.

[130] Narae Shin, Seung Hwan Lee, Viet Anh Pham Ba, Tai Hyun Park, and Seunghun Hong. Micelle-stabilized olfactory receptors for a bioelectronic nose detecting butter flavors in real fermented alcoholic beverages. *Scientific reports*, 10(1):9064, 2020.

[131] Satpreet H. Singh, Floris van Breugel, Rajesh P. N. Rao, and Bingni W. Brunton. Emergent behaviour and neural dynamics in artificial agents tracking odour plumes. *Nature Machine Intelligence*, 5(1):58–70, Jan 2023.

[132] Jake Snell, Kevin Swersky, and Richard S. Zemel. Prototypical networks for few-shot learning, 2017.

[133] Stanley S. Stevens. *Handbook of experimental psychology*. Wiley, 1951.

[134] Haiyue Sun, Shunda Qiao, Ying He, Xiaorong Sun, and Yufei Ma. Parts-per-quadrillion level gas molecule detection: Co-lites sensing. *Light: Science & Applications*, 14(1):180, Apr 2025.

[135] Seung-Hyun Sung, Jun Min Suh, Yun Ji Hwang, Ho Won Jang, Jeon Gue Park, and Seong Chan Jun. Data-centric artificial olfactory system based on the eigengraph. *Nature Communications*, 15(1):1211, Feb 2024.

[136] Ilya Sutskever. Neurips test of time talk - sequence to sequence learning with neural networks: what a decade. YouTube, December 2024. Uploaded by seremot on 2024-12-13, URL: `https://www.youtube.com/watch?v=1yvBqasHLZs`.

[137] Arvin Tashakori, Zenan Jiang, Amir Servati, Saeid Soltanian, Harishkumar Narayana, Katherine Le, Caroline Nakayama, Chieh-ling Yang, Z. Jane Wang, Janice J. Eng, and Peyman Servati. Capturing complex hand movements and object interactions using machine learning-powered stretchable smart textile gloves. *Nature Machine Intelligence*, 6(1):106–118, Jan 2024.

[138] Thomas Tille. Automotive suitability of air quality gas sensors. *Sensors and Actuators B: Chemical*, 170:40–44, 2012. Eurosensors XXIV, 2010.

[139] Luca Turin, Simon Gane, Dimitris Georganakis, Klio Maniati, and Efthimios M. C. Skoulakis. Plausibility of the vibrational theory of olfaction. *Proceedings of the National Academy of Sciences*, 112(25):E3154–E3154, 2015.

[140] Alexander Vergara. Gas sensor arrays in open sampling settings. UCI Machine Learning Repository, 2013. DOI: https://doi.org/10.24432/C5JP5N.

[141] Matt Wachowiak, Adam Dewan, Thomas Bozza, Tom F. O'Connell, and Elizabeth J. Hong. Recalibrating Olfactory Neuroscience to the Range of Naturally Occurring Odor Concentrations. *The Journal of Neuroscience*, 45(10):e1872242024, March 2025.

[142] Alex Wang, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel R. Bowman. Glue: A multi-task benchmark and analysis platform for natural language understanding, 2019.

[143] Guangzhi Wang, Yuchen Guo, Yang Yu, Yan Shi, Yuxiang Ying, and Hong Men. Colornet: An ai-based framework for pork freshness detection using a colorimetric sensor array. *Food Chemistry*, 471:142794, 2025.

[144] Zhenhailong Wang, Manling Li, Ruochen Xu, Luowei Zhou, Jie Lei, Xudong Lin, Shuohang Wang, Ziyi Yang, Chenguang Zhu, Derek Hoiem, Shih-Fu Chang, Mohit Bansal, and Heng Ji. Language models with image descriptors are strong few-shot video-language learners. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22, Red Hook, NY, USA, 2022. Curran Associates Inc.

[145] Jennifer N. Wei, Carlos Ruiz, Marnix Vlot, Benjamin Sanchez-Lengeling, Brian K. Lee, Luuk Berning, Martijn W. Vos, Rob W.M. Henderson, Wesley W. Qian, D. Michael Ando, Kurt M. Groetsch, Richard C. Gerkin, Alexander B. Wiltschko, Jeffrey Riffel, and Koen J. Dechering. A deep learning and digital archaeology approach for mosquito repellent discovery. *bioRxiv*, 2024.

[146] Laura Weidinger, Jonathan Uesato, Maribeth Rauh, Christopher Griffin, John Mellor, Po-Sen Huang, Amelia Glaese, Borja Balle, Atoosa Kasirzadeh, Iason Gabriel, et al. Taxonomy of risks posed by foundation models. *arXiv preprint arXiv:2208.05300*, 2022.

[147] Geemi P. Wellawatte and Philippe Schwaller. Human interpretable structure-property relationships in chemistry using explainable machine learning and large language models. *Communications Chemistry*, 8(1):11, Jan 2025.

[148] Kathleen E Whitlock. Genetic variation in human odor perception. *Chemical Senses*, 45:549–556, 2020.

[149] Wikipedia contributors. Optimus (robot) — wikipedia, the free encyclopedia, n.d. Accessed: 2025-05-23.

[150] CJ Woolf, L Ma, S Qian, YB Sirotin, and D Rinberg. Detection and discrimination of low-concentration odorants in complex mixtures. *Current Biology*, 32(1):89–98, 2022.

[151] Junbin Xiao, Xindi Shang, Angela Yao, and Tat-Seng Chua. Next-qa:next phase of question-answering to explaining temporal actions, 2021.

[152] Fengli Xu, Qianyue Hao, Zefang Zong, Jingwei Wang, Yunke Zhang, Jingyi Wang, Xiaochong Lan, Jiahui Gong, Tianjian Ouyang, Fanjin Meng, Chenyang Shao, Yuwei Yan, Qinglong Yang, Yiwen Song, Sijian Ren, Xinyuan Hu, Yu Li, Jie Feng, Chen Gao, and Yong Li. Towards large reasoning models: A survey of reinforced reasoning with large language models, 2025.

[153] Jun Xu, Tao Mei, Ting Yao, and Yong Rui. Msr-vtt: A large video description dataset for bridging video and language. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.

[154] Eugene Yee, R. Chan, P. R. Kosteniuk, G. M. Chandler, C. A. Biltoft, and J. F. Bowers. Measurements of level-crossing statistics of concentration fluctuations in plumes dispersing in the atmospheric surface layer. *Boundary-Layer Meteorology*, 73(1-2):53–90, 1995.

[155] Xiaomeng Yin, Hao Zhang, Xuezhi Qiao, Xinyuan Zhou, Zhenjie Xue, Xiangyu Chen, Haochen Ye, Cancan Li, Zhe Tang, Kailin Zhang, and Tie Wang. Artificial olfactory memory system based on conductive metal-organic frameworks. *Nature Communications*, 15(1):8409, Sep 2024.

[156] Xueying Zhan, Qingzhong Wang, Kuan hao Huang, Haoyi Xiong, Dejing Dou, and Antoni B. Chan. A comparative survey of deep active learning, 2022.

[157] Jieyu Zheng and Markus Meister. The unbearable slowness of being: Why do we live at 10 bits/s? *Neuron*, 113(2):192–204, Jan 2025.