**Literature Review:**

[**“Gemini Robotics: Bringing AI into the Physical World”**](https://arxiv.org/pdf/2503.20020)  
**Authors**: Gemini Robotics Team, Google DeepMind  
**Year**: 2025

**Overview**

This paper presents *Gemini Robotics*, a new family of AI models purpose-built for robotics, developed by Google DeepMind. Building on the Gemini 2.0 foundation, the models are designed to enable general-purpose, safe, and dexterous robotic control by integrating vision, language, and action (VLA) capabilities.

**Key Contributions**

1. **Gemini Robotics-ER (Embodied Reasoning)**
   * A vision-language model (VLM) enhanced for physical world understanding.
   * Capable of 2D/3D object detection, pointing, trajectory prediction, and grasp planning.
   * Introduces **ERQA**, a novel benchmark to evaluate embodied reasoning.
   * Trained on large-scale spatial and temporal datasets, allowing high-level perception and reasoning.
2. **Gemini Robotics (VLA Model)**
   * A full vision-language-action model enabling **direct robotic control**.
   * Learns dexterous tasks (e.g., origami folding, card games) with smooth, reactive movements.
   * Capable of zero-shot and few-shot learning, supporting rapid adaptation to novel tasks or robot types.
3. **Adaptation and Specialization**
   * Gemini Robotics can fine-tune with only ~100 demos for new short-horizon tasks.
   * Demonstrates transfer to **new embodiments**, such as bi-arm or humanoid robots.
   * Capable of advanced physical interaction under open-vocabulary language commands.
4. **Responsible AI in Robotics**
   * Includes discussion on safety, adversarial robustness, and ethical deployment.
   * Developed under Google’s AI Principles with an emphasis on societal impact mitigation.

**Technical Strengths**

* **Multimodal Fusion**: Integrates vision, language, and robot control within a unified architecture, outperforming prior single-modality or narrow-domain robotics models.
* **Generalization**: Demonstrates cross-task, cross-domain, and cross-embodiment generalization with high reliability, even under instruction variations or novel object positions.
* **Speed and Dexterity**: Designed for low-latency control, enabling precise and fluid manipulation.

**Limitations**

* **Spatial-Temporal Grounding**: Gemini 2.0 and its successors struggle with long video contexts and precise numerical outputs (e.g., bounding box precision).
* **Simulation to Reality Gap**: More work is needed to improve sim2real transfer and robustness in highly contact-rich environments.
* **Multi-step Reasoning**: Performance degrades on complex tasks requiring long-horizon planning and execution.

[**Trust and Acceptance of AI Caregiving Robots: The Role of Ethics and Self-Efficacy**](https://pdf.sciencedirectassets.com/783247/1-s2.0-S2949882124X00031/1-s2.0-S2949882124000756/main.pdf?X-Amz-Security-Token=IQoJb3JpZ2luX2VjEGwaCXVzLWVhc3QtMSJGMEQCIEoIzquA7ChJCjhP%2B7y5jTm4nJTFZAymab9BldF%2FpqEAAiA0iZYF%2FwaHicCjWiRtMFdqNczr%2BZI0X7CNukJqZMWPvCqyBQhUEAUaDDA1OTAwMzU0Njg2NSIMqmWnROCBfkqYiT8iKo8FkaLt7cnVnVzELeycI9fIxGLW0Xu2OzYuF1cVn7QMQpBSAD3YGErF8sIR6lDjFCocS%2BuidHYgFk%2Foz9rDQp%2FCnichlTjY%2ByU%2BXWnoRHBvjz4rPKVhsjwfPzK9HLjHluL0bD6MVP61WdGRe9jeD%2FQy69nj%2FrvrF0pBkrmR6EkXDvAp8Y8vGMFAHtvtdEh%2Fq%2F8kcHXfL4PdCUnMjJJ5M0ZrQmNTzt1nLVxAlAwwDy%2B1AiktF6zGE3Bj6kSavqxMdDeehhoz3dCPi0z7RtEXl%2Fo8UL1BCKYcX0p57AGA59FeFwWSJH15HbMQn7pMS0fJevTwEIMTBYW0i%2ByBii2ogOdPEDoWxidPWgLcQ0pE7dyH9doIyQXWsKtnI3d96ZOWAal08VFUep4H%2BWQgTupgTfek1CAaytWioX0nSpnUViC2N19mDZZBvHbyxcVBf7Tl39kPWW13fsuVbfWyl3a7zLkqpUBLXypmFhM9wy%2FHOIraLbdsEc7ViOZ%2BMLhkHpsRpvFfpSjd9memUwBNqWaRz4ePgdAMK98Q76AmQPMRXX7Lr2GXO8mFNe35H1ZyTEKRjFxEjNGLGF7WVsWjPDyr%2BFL%2B0DbF%2FBbMEdArTvgWSuWoxT%2BMUC25Zwp21cjBiMMqBKr4uW7AFDHpUcz2laUykHSfBiuzoD5qmigsBQ6H0OMJQY95tHmIqp4zq29dtS0io%2Fl7YaBEmCFELvGf3LPtxZyQAN%2FcUUbLtQZRq6%2F2RloDOg7%2BQQv6bLRhfn083zwZ8FXZjqbjS4gCU4WfQnhxJu97auRpD91NmiQLGQny8ZnZ%2BSid9EgDEg02rdyi2SOfbhwZKY0Gvc677S8AnR4MaF242VC93CormD71yHJsODY9hDDxlL7CBjqyAcvkHnckIOqEijiXNYYo7SIkVNIOcVRDx1aFDDZomZg6cbeffAN5AEBv7xiMJ7vmBM%2FisWVq1tPSeN7d7Mc%2BT1GYDWEPO8Ng4cl%2F3ogQOZ%2BNALxmDl5ritE1aI4h6w4xlVuqSuXpWmEyEjQjjP2pdcoWN9sVqOunfDRobtt3%2BF0AbJjn3lgXLVcCm0yLgIsXo%2FZ1sj65jspnjTdk4G%2FFTvOYpAKXuS%2FdJk1%2FProiHfP6fss%3D&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Date=20250616T034706Z&X-Amz-SignedHeaders=host&X-Amz-Expires=300&X-Amz-Credential=ASIAQ3PHCVTY5A7NTMNZ%2F20250616%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Signature=5fa8ec1786293d6b308ea0cd8a2b4a1953e46f97ce157971847377dde5e80aa9&hash=77ab041ebb5bd77ef7886680ada2e3a47596d8c3991499d3c85c5ba7b7f4094b&host=68042c943591013ac2b2430a89b270f6af2c76d8dfd086a07176afe7c76c2c61&pii=S2949882124000756&tid=spdf-38d06903-9360-4657-92f7-2c125691aa3b&sid=85ecd19e7d297341356b6d0736a1840a1ab8gxrqa&type=client&tsoh=d3d3LnNjaWVuY2VkaXJlY3QuY29t&rh=d3d3LnNjaWVuY2VkaXJlY3QuY29t&ua=0f155b5b5f53555506&rr=950752ec5a51699c&cc=us)

**Authors**: Cathy S. Lin, Ying-Feng Kuo (corresponding author), Ting-Yu Wang  
**Affiliation**: Department of Information Management, National University of Kaohsiung, Taiwan  
**Year**: 2025

**Overview**

This study examines the influence of ethical concerns and self-efficacy on the public’s trust and willingness to use AI caregiving robots, particularly in healthcare and eldercare. By developing a new AI ethics scale and expanding the Unified Theory of Acceptance and Use of Technology (UTAUT), the research provides insights into factors that increase acceptance of caregiving robots.

**Key Contributions**

1. **AI Ethics Scale (Second-Order Construct)**
   * Developed and validated across multiple user groups.
   * Composed of three key dimensions:
     + **Transparency** – clarity in data use and AI functioning.
     + **Accountability** – responsibility mechanisms when AI causes harm.
     + **Fairness** – avoidance of algorithmic bias and discrimination.
2. **Expanded UTAUT Framework**
   * Integrates ethics and trust, often omitted in classic models.
   * Includes: social influence, performance expectancy, self-efficacy, and trust.
   * Demonstrates:
     + Ethics and social influence increase self-efficacy.
     + Self-efficacy strengthens trust and performance expectations.
     + Trust and performance expectancy directly increase intention to use AI caregiving robots.

**Methodology**

* **Participants**: 391 survey respondents, including both potential users and future developers.
* **Analysis**: Partial Least Squares Structural Equation Modeling (PLS-SEM).
* **Validation**: Strong reliability across gender and role subgroups (users vs. developers).

**Key Findings**

* **Trust** is the strongest direct predictor of intention to use caregiving robots.
* Adding **AI ethics** significantly enhances the model's explanatory power.
* **Developers** are more influenced by ethics; **users** by social influence.
* **Gender differences**:
  + Men respond more to ethical concerns.
  + Women are more influenced by social norms.

**Practical Implications**

* **For Developers**: Design AI systems with embedded ethical principles.
* **For Organizations**: Clearly explain how AI works, including data handling and limitations.
* **For Policymakers**: Implement comprehensive ethical frameworks for caregiving AI.
* **For Users**: Knowing that ethics are prioritized helps build trust and adoption.

**Limitations and Future Work**

* **Privacy** excluded due to external regulations (e.g., GDPR).
* **Cross-sectional design** limits temporal insights: future studies could use qualitative or longitudinal approaches for deeper understanding.

[**Principles and Guidelines for Evaluating Social Robot Navigation Algorithms**](https://dl.acm.org/doi/pdf/10.1145/3700599)

**Authors**: A. Francis, et al.  
**Published in**: *ACM Transactions on Human-Robot Interaction*  
**Year:** February 2025

**Overview**

This paper addresses the lack of standardized evaluation frameworks in **social robot navigation**, where robots must operate in dynamic, human-centered environments. The authors propose a taxonomy of key principles, metrics, and evaluation guidelines to enable **repeatable, comparable, and meaningful benchmarking** across different platforms and research groups.

**Key Contributions**

1. **Definition of Social Navigation**
   * A socially navigating robot is one that adheres to principles such as:
     + **Safety**, **Comfort**, **Legibility**, **Politeness**,
     + **Social Competency**, **Proactivity**, **Context Awareness**, and **Responsiveness**.
2. **Evaluation Framework**
   * Introduces a **metrics framework** that includes both subjective (human-rated) and analytic (computable) evaluations.
   * Encourages use of **common APIs and scenarios** to ensure cross-platform comparability.
3. **Guidelines for Benchmarking**
   * Details best practices for designing **scenarios, datasets, and simulators**.
   * Proposes **recommendations for metric usage**, including success, collisions, social compliance, and comfort-related measures.

**Methodology & Tools**

* **Taxonomy**: Developed through expert discussions at the Social Navigation Symposium.
* **Scenarios and Simulators**: Standardized design guidelines for social navigation tasks.
* **Metrics**:
  + **Hand-crafted**: Success Rate, Collision Rate, Time to Goal, Path Length, Space Compliance.
  + **Subjective**: Human-rated surveys focusing on safety, comfort, and perceived intent.
  + **Learned**: Ongoing efforts to train ML models on large-scale survey-labeled datasets.

**Key Findings**

* **Lack of standardization** is a key barrier to fair comparison in social navigation research.
* **No single metric** can capture all desired behaviors—**context and task specificity** matter.
* **Human surveys** are gold standard but expensive and inconsistent unless validated properly.
* **Learned metrics** (e.g., time-to-collision, comfort indices) offer potential but lack maturity.
* **Clear scenario definitions and reporting standards** are essential for reproducibility.

**Practical Implications**

* **Researchers**: Can build on a common framework for evaluating social navigation.
* **Developers**: Encouraged to design navigation algorithms with **social acceptability** in mind.
* **Benchmarking bodies**: Should work toward shared APIs and simulation interfaces.
* **Human-Robot Interaction (HRI)** studies: Need better standardization in survey design and deployment.

**Limitations and Future Directions**

* **Survey fatigue** and subjective variability pose challenges in large-scale evaluations.
* **No current learned metric** fully substitutes for human judgments across all tasks.
* Future work should focus on:
  + Creating **large, annotated datasets** with both human and algorithmic metrics.
  + Developing **generalizable learned proxies** for surveyed feedback.
  + Improving cross-cultural and multi-stakeholder benchmarking.