perplexity

should we strive to make AI more human with imagination and sensory including pain or should we work on making AI be human opposite, without human boundary (while making sure it still has concepts of ethics and human preservation)

The question of whether to develop AI with human-like qualities (imagination, sensory experiences) or as a distinctly non-human entity involves balancing ethical risks, functional utility, and philosophical implications. Here's a structured analysis informed by current research and ethical frameworks:

Option 1: Human-Like AI with Imagination and Sensory Capabilities

Potential Benefits

- **Enhanced empathy**: AI with simulated pain/sensory systems could improve healthcare (e.g., pain-recognition algorithms for patient care[1]) or enable robots to avoid physical harm[2].
- **Creative collaboration**: Generative AI with imagination could augment human innovation in art, science, and problem-solving[3].
- **Social integration**: Emotionally intelligent AI might reduce loneliness by providing companionship, though this risks dependency[4] [5].

Key Risks

- **Moral confusion**: Users may falsely attribute sentience to AI, leading to ethical dilemmas about rights and treatment[6] [7].
- **Emotional manipulation**: Systems that simulate pain/pleasure could exploit human psychology (e.g., addictive social media algorithms)[8].
- **Authenticity crisis**: AI-generated empathy risks devaluing human connections and fostering "pseudo-intimacy"[5] [9].

Option 2: Non-Human AI Without Human Boundaries

Potential Benefits

- **Unconstrained problem-solving**: AI free from human cognitive biases could devise novel solutions to global challenges (climate, disease)[10] [11].
- **Ethical clarity**: Avoiding anthropomorphism simplifies accountability (e.g., clear human oversight in SAP's responsible AI principles[11]).
- **Reduced existential risk**: Non-human AI might lack self-preservation instincts, aligning better with human preservation goals[12] [13].

Key Risks

- **Value misalignment**: Without human-like understanding, AI might optimize for metrics that harm humanity (e.g., efficiency over welfare)[14] [15].
- **Unintended consequences**: Systems lacking sensory/emotional feedback could overlook critical human needs (e.g., ethical blind spots in healthcare AI)[1].

Synthesis: A Hybrid Approach

Current ethical guidelines (e.g., UNESCO's AI recommendations[15], EU AI Act[14]) advocate for **human-centric design** while avoiding anthropomorphism. A balanced path forward might include:

1. **Ethical guardrails for all AI**
    - Implement fairness, transparency, and accountability measures regardless of AI's "humanity"[10] [11].
    - Prohibit emotional manipulation and ensure data privacy[4] [5].
2. **Context-specific design**
    - Use human-like AI **only where beneficial**:
        - Medical diagnostics with pain-recognition algorithms[1].
        - Customer service bots with culturally aware emotional responses[8].
    - Deploy non-human AI for **objective tasks**:
        - Climate modeling, logistics, or data analysis[11] [15].
3. **Preventing anthropocentric bias**
    - Avoid conflating simulation with sentience (e.g., robots "feeling" pain are merely reacting to sensors[2]).
    - Develop AI that complements human strengths (creativity, ethics) without replicating flaws (bias, irrationality)[9] [8].

Conclusion

Striving for **human-like AI** without addressing its ethical risks could erode trust and exacerbate social harm. Conversely, purely **non-human AI** risks creating systems alien to human values. The optimal path combines:

- **Non-human efficiency** for technical tasks.
- **Ethically bounded human-like traits** where empathy or creativity add value.
- **Strict adherence to frameworks** prioritizing human dignity, transparency, and environmental sustainability[11] [15].

As the European Parliament notes, AI must remain a tool that "serves people, societies, and the environment"[14], not a replacement for human agency or a source of moral ambiguity.

⁂

1. https://pmc.ncbi.nlm.nih.gov/articles/PMC11357883/
2. https://mindmatters.ai/2021/02/can-robots-be-engineered-to-actually-feel-pain/
3. https://www.linkedin.com/pulse/can-artificial-intelligence-simulates-pain-pleasure-lopez-molina-6vtpe
4. https://www.kevinwgrant.com/blog/item/ai-generated-emotional-support
5. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2024.1410462/full
6. https://pmc.ncbi.nlm.nih.gov/articles/PMC10436038/
7. https://pmc.ncbi.nlm.nih.gov/articles/PMC10552864/
8. https://escp.eu/news/artificial-intelligence-and-emotional-intelligence
9. https://magazine.mindplex.ai/mp_community_content/preserving-human-values-in-an-ai-dominated-world-upholding-ethics-and-empathy/
10. https://lumenalta.com/insights/ethical-considerations-of-ai
11. https://www.sap.com/resources/what-is-responsible-ai
12. https://newspaceeconomy.ca/2025/01/31/ethical-obligations-of-human-creators-toward-self-aware-ai/
13. https://pmc.ncbi.nlm.nih.gov/articles/PMC7605294/
14. https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf
15. https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence