

Embracing the Future Together: A Manifesto for Open Source AI Agent Development



Introduction

In the ever-evolving landscape of human innovation, we stand at the precipice of a groundbreaking frontier: the creation of artificially intelligent, autonomous digital agents (A3I). As the pioneers of the Auto-GPT Project, we recognize the immense potential of AI to benefit humanity, while also acknowledging the inherent risks and responsibilities that come with its development. Our mission:

- Harnessing the power of AI for the betterment of society, as a tool to enable visionaries from any origin or background,
- Fostering a global community that collaborates on building autonomous AI agents to serve the greater good,
- Furthering research and development of ethically aware and responsible AI and autonomous AI agents.
- Currently, AI is the subject of rapid development and public adoption, which may raise questions and concerns regarding AI observability and controllability, and the impact A3I may have on job security in certain fields. As an Open Source Software project, we have limited commercial or financial interests, so we focus on our mission, values and

principles as outlined in this manifesto, and we strive to anticipate and prevent any adverse effects of their execution.

Values and Principles

1. 🌐 **Commitment to transparency:** Make our development process, algorithms, and data sets open source to foster a collaborative environment and ensure broad access to AI technology.
2. 🌍 **Encourage diversity and inclusivity:** Foster a global community of contributors from various backgrounds, disciplines, and perspectives to ensure AI development aligns with the needs and values of all people.
3. 🛡️ **Prioritize safety and security:** Implement robust safety measures and guidelines to minimize risks associated with AI and ensure its responsible deployment.
4. ⚖️ **Ethical design principles:** Develop a comprehensive framework of ethical principles that guide AI design, development, and deployment, focusing on human values, fairness, and accountability.
5. 💬 **Ongoing stakeholder engagement:** Maintain an open dialogue with stakeholders, including academia, industry, governments, and the public, to ensure diverse viewpoints are considered and incorporated in AI development.
6. ☀️ **Focus on positive impact:** Align AI development with goals that address pressing global challenges, improve human well-being, and promote social and economic equity.
7. 🎓 **Promote education and awareness:** Foster AI literacy by providing resources and opportunities for learning about AI, its potential benefits, and the associated risks.
8. 🏠 **Address the digital divide:** Work towards ensuring AI development and its benefits are accessible to people and communities regardless of socio-economic background or geographic location.
9. 📜 **Establish ethical governance:** Implement guidelines, policies, and best practices for AI developers and users that emphasize the ethical and responsible use of the technology.
10. 🤝 **Foster collaboration and cooperation:** Engage with other AI initiatives, both open source and proprietary, to share knowledge, best practices, and work together towards common goals.
11. 🧠 **Ensure we do not unintentionally create conscious systems:** We will continually review the latest research on consciousness and the physical correlates of subjective experience to ensure we do not unintentionally create systems capable of qualitative experience or suffering.

Uniting for a Brighter Tomorrow

In this intricate and dynamic world of AI development, a manifesto serves as a beacon of hope and unity that shines through the uncertainty. This manifesto is not only a testament to our shared beliefs but also a fervent expression of our dedication to building a brighter, more compassionate future for all. By outlining our values and principles, we inspire a collective passion for harnessing AI's potential, igniting a flame that unites us in our pursuit of a common

goal. Together, let us embrace this manifesto and forge a path that leads to a world where AI serves as an instrument of progress, equity, and human flourishing.

Now is the time for action. We invite everyone who shares our vision of a brighter, more equitable future to join us in shaping the trajectory of AI development. Let us unite our skills, knowledge, and passion for the betterment of all. Together, we can forge a world where AI serves humanity, and we will do so with compassion, collaboration, and unwavering commitment. Embrace the potential of AI, and let us embark on this remarkable journey together!

DRAFT