# DEVELOPING AN AI ALIGNMENT REGIMEN FOR SUPPORTING PRODUCTION ENTREPRENEURS

Mihai Dragomir[1], Daniela Popescu[1], Ștefania Chiribău-Vitlinger[1] and Diana Dragomir[1]
[1]Technical University of Cluj-Napoca, Cluj-Napoca, Romania
Corresponding author: ORCID No. 0000-0003-0561-4274, mihai.dragomir@muri.utcluj.ro

ABSTRACT: This paper analyzes the need to infuse human-centric values characteristic of the management style of production entrepreneurs into the AI instruments to be used in industrial companies. A careful detailing of the human related elements, the sustainability related functions and the technological features is performed in relation to the need to customize the training of AI systems that are employed in production processes, along the value creation chain. This is then compared with the results of a focus group of specialists in the field of production in order to generate an action plan and regimen to support AI alignment with minimal disruption to the companies' results and performance.
KEYWORDS: AI alignment, production entrepreneurs, production processes, human-centric values

## 1. INTRODUCTION

The use of AI in various industries is moving slowly from the trendy phase into the must-have phase and production companies are no exception to this transformation. This is even more true in the case of entrepreneurs starting or significantly overhauling a manufacturing enterprise. The media and most regular users are focused on Generative AI like Chat-GPT, Gemini or Copilot, or their counterparts in creating AI art or AI videos, but in essence a production environment is suitable for employing machine learning based on IoT sensors and cyber-physical devices. If the first category of tools mentioned can be leveraged for improving customer engagement and requirements engineering, the machine learning tools, and the decision support systems can strongly impact on the actual manufacturing environment in terms of interactions between humans and machines and prediction and prevention of costly quality, safety or maintenance issues.

In all cases, ensuring the AI systems are fully complementary with the human elements, which is prevalent especially in start-ups, can be achieved by properly aligning the AI to the values and action principles that make up the personality of the founders and leaders of the company. The current paper proposes a methodological approach to achieving this alignment through a combination of logical approaches and co-creation demarches that aim to capture the managerial style and intent.

## 2. LITERATURE REVIEW

Mapping the research landscape of AI alignment for entrepreneurship is a difficult task, as it involves a multitude of facets. For example, the authors [1] focus on the linguistic aspects that showcase the intrinsic values and the ways in which AI systems such as LLM can adopt human-like patterns. In another paper [2] discusses trust level in the AI-human relation from the point of view of the decision to include certain values in the behavior and to establish the level to which the system follows these values to mimic humans. In a more technical approach, [3] emphasize the need to advance mutual trust by ensuring security of data, facilitating the relation with quantum computing and ensuring decentralized performance in machine learning. At the same time, [4] analyzes the negative effects of the widespread and poorly aligned recommender systems in various types of applications and underlines their many complexities.

Entrepreneurship is believed to be significantly impacted by the widespread use of digital technologies, with AI occupying a central position within the research landscape investigated by researchers [5]. The intricate relationship of AI and entrepreneurship is heavily influenced by the ability of future endeavors in either business or research to involve the humans and the automated systems in synchronicity [6]. Nonetheless, since AI adoption can have important negative consequences especially on small business [7], a rush to implement solutions with no regard for ethical, moral and humanistic considerations will most probably not lead to the desired outcomes of market validation and financial success.

## 3. METHODOLOGY AND RESULTS

In order to carry out the research work for the current paper, the following research methodology has been implemented:

1. Identifying challenges and goals of an entrepreneurship Al support system.
2. Processing useful Industry 5.0 concepts, Al strategies and action plans.
3. Brainstorming focus group with 5 production engineers & managers.
4. Validation with a second focus group of 5 production entrepreneurs.

The general concept employed started from the three main alignment direction identified based on literature and production sector trends (Figure 1): instilling human-centric values, tackling the sustainability challenges and making mindful and responsible use of advanced technologies. The following chapter explains the results of implementing this approach, focusing on interpreting and drawing workable conclusions from the interactions with the two focus groups. The presentation is performed in accordance with the needed inputs (requirements) that AI systems' developers use.

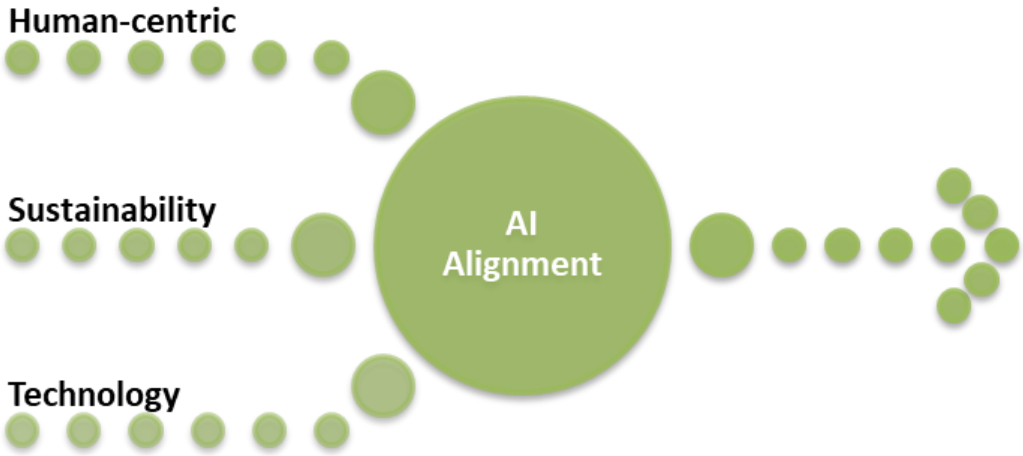


**Figure 1.** Overall AI alignment concept for production entrepreneurship

### 3.1 AI to support technology and vice versa

Use Al platforms and solutions (generative Al, Large Language Models, machine learning, deep learning) in an industrial medium characterized by the abundance of data provided by lot sensors becomes a necessity under the strict competition conditions of today. The tasks of sensor data fusion, pattern detection and decision support need to be quickly solved for middle and top managers, and for automated decision systems, in order to allow them to detect and remove any negative impact produced by product scrap, process discharge and emissions or product life-cycle utilization

An approach with 4 stages is desired by the industry representatives when presented with the current state of the art: (i) Architecture development, (ii) Manufacturing data collection including the installation of IoT sensors and growing the necessary database (iii) Production tools adaptation to the needs of Industry 5.0 performance management, and (iv) System development/customization, testing and deployment under real production conditions.

The first focus group addressed two high-level manufacturing domains (the automotive and aerospace industries), as well as two consumers goods related domains (the furniture and apparel industries) and one classic metal fabrications company. Training the Al models in trend detection, uncertainty measurement and synergic effects is the critical element identified following the discussions. Allowing the users access to ready-made solutions that not only process and monitor performance but also permit them to quickly crowd-source improvement ideas and track their implementation is seen as a key driver of innovation in this area.

The database should be extendable with further information as it becomes available and is collected, including information resulting from the application of previous recommended measures, thus enabling the charts, graphs and indicators to exhibit an emergent behavior in which their interpretation becomes more complex over time and permits the users to pilot a deep production system augmentation without incurring extensive costs. Production engineers tend to be focused on process related issues like uptime, overall equipment effectiveness and scrap processing, while the production managers are mostly interested in ensuring the integrity of the supply chain and a proper cost-results balance with respect to the necessary equipment and software investments.

### 3.2 AI to support sustainability - SDGs

The impact of Al instruments is mostly visible in the studied situation on the Sustainable Development Goal SDG 13 - Climate change Action. This support should be implemented in two dimensions:

- current contributions (especially in the past 2 years) related to the climate performance of the manufacturing sector, and
- future potential on enhancing the strategy development and implementation process in the same sector (especially related acceptance and compliance).

The first aspect will target the degree of use and perceived helpfulness of using LLM, Al art generators and ML techniques to craft and disseminate company and sectoral strategies and instruments about climate concerns (e.g., carbon management for products and processes, GHG emissions, preparedness for the negative effects of climate change, raising the awareness of customers, etc.). The second aspect will be investigated through simulations of possible uses of the Al tools in delivering better strategies and plans and a more effective implementation of these, targeting issues related to creating meaningful awareness among customers by improved visual communication, writing better documents based on real data from industry and the markets, and uncovering trends, patterns, non-obvious connections and emergent system behaviors.

The relationship with the companies and the customers will be managed through meetings to take place online and offline, aided by specific AI platforms. Once collected, the qualitative data will be processed using excellence and maturity frameworks such as CMMI and MBNQA/EFQM to determine the capability of existing strategic documents and initiatives to determine positive changes. When investigating the future potential of Al tools, forecasting and foresight analytical methods such as brainstorming and mind-maps should be for determining the implementation directions, FMEA - Failure Modes Effects and Analysis and What-if scenarios are needed for limiting the possible negative correlations, while multi-criterial decision analysis for are needed validating the obtained findings against the perception of firms and interested people.

Combining these methods in innovative ways can generate the understanding of complex situations, which is in line with the intricacy associated with climate actions, whence iterative improvement is desirable and effective in the long run. However, in this case, as opposed to the previous one where there was an agreement between the two focus groups, the first meeting was only partially validated by the second group, as production entrepreneurs expressed concerns that an incipient preoccupation with climate actions and the implementation of complex tools, will make their efforts difficult in initiating a new business. Although, when asked about the possibility of applying the approach in a second-chance scenario, with a new venture after one or more failed ones, this objection was removed.

### 3.3 AI to support humans and human-aligned AI

Having Al support human-centricity of a production sector approach can seem paradoxically, as Al tools are under scrutiny for their capability in replacing humans in many fields. However, within the Industry 5.0 concept, humans take a more center stage role and are valued for their unique contributions, while Al is expected to automate and pre-solve many low stakes but high work volume issues.

The involvement of human in manufacturing is difficult to replace in product development, customer experience and even marketing. On the other hand, Al can contribute to safer workplaces by data processing and trend detection of hazards and harmful events, while at the same time being an excellent tool in optimizing process performance, inventory planning, inside and outside logistics, distribution networks, customer service history and many other data heavy aspects.

The contributions of Al in ensuring predictive maintenance, proactive quality assurance and lean management are also well documented and should form part of the planned alignment regimen. Moreover, when training AI systems, especially the generative tools and the image processing, properly documenting the experience, emotional responses and even idiosyncrasies (e.g., intuition) of human decision makers or human operator has the potential to enhance the integrated AI approach. In this case, the production entrepreneurs' focus group was enthusiastically more ready to discover and apply new tools and models, in comparison with the seasoned first group composed of experienced personnel from production companies.

## 4. ENTREPRENEURIAL INSIGHT

To complete the validation process, the second focus group debated also the LLM answer to a specific question about supporting production entrepreneurs (Figure 2).

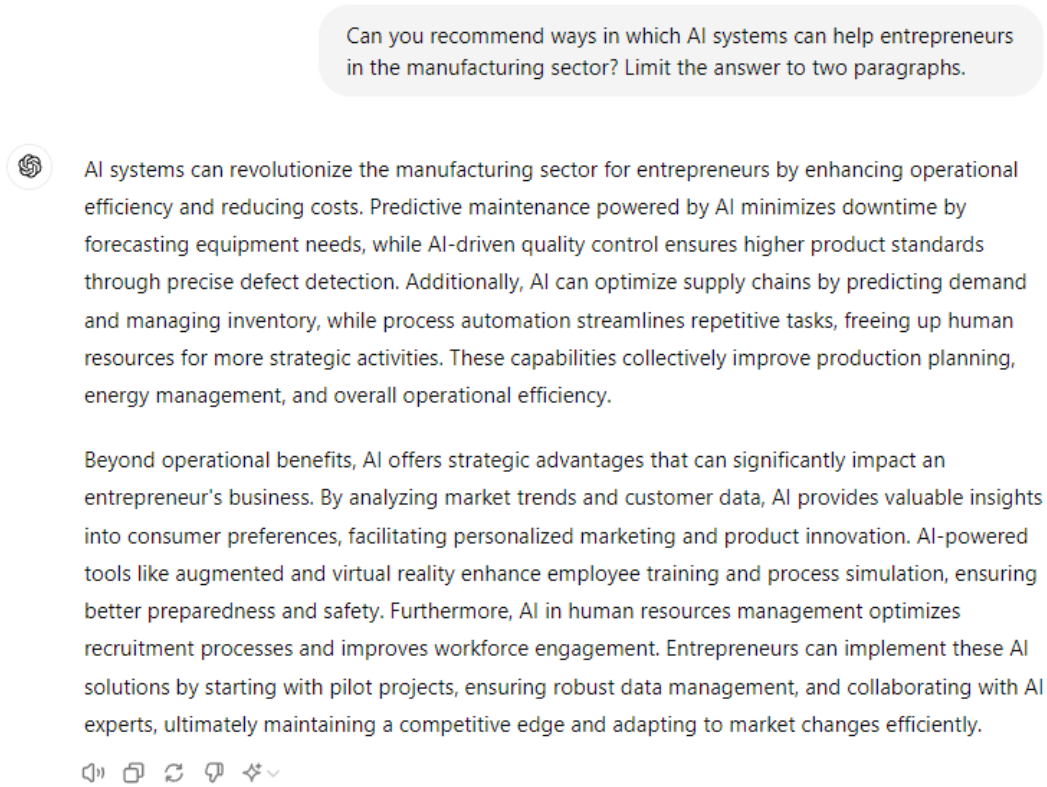


**Figure 2.** AI LLM generated answer for entrepreneurs. Source: Open AI ChatGPT, https://chatgpt.com/

All the proposed process improvement measures have been favorably evaluated by the group, although further details and measures were required. Starting with small pilots is also seen as the only way they would begin to approach AI for their own start-up ventures. Still, employing AI for human resources recruitment and participation is considered risky in the case of newly funded companies.

Overall, the degree of agreement, factoring in also the local regulatory framework, can be appreciated at ca. 50% between the real production entrepreneurs and the ChatGPT provided answer that analyzes their situation.

## 5. CONCLUSIONS

In conclusion, this work has the role of pinpointing the main issues in developing the proper ways to align AI to the values of production entrepreneurs, with further in-depth research needed for a complete understanding and the development of proper measures. However, the 3 main directions identified and the way to approach each of them can constitute the roots of an AI alignment regimen applicable to generative AI solutions.

## 6. REFERENCES

[1] A. Kasirzadeh and I. Gabriel, "In Conversation with Artificial Intelligence: Aligning language Models with Human Values," *Philosophy & Technology,* vol. 36, p. 27, 2023.

[2] L. Bojic, "AI alignment: Assessing the global impact of recommender systems," *Futures,* vol. 160, p. 103383, 2024.

[3] A. Habbal, M. Khalif Ali and M. Ali Abuzaraida, "Artificial Intelligence Trust, Risk and Security Management (AI TRiSM): Frameworks, applications, challenges and future research directions," *Expert Systems with Applications,* vol. 240, p. 122442, 2024.

[4] D. A. Shepherd and A. Majchrzak, "Machines augmenting entrepreneurs: Opportunities (and threats) at the Nexus of artificial intelligence and entrepreneurship," *Journal of Business Venturing,* vol. 37, p. 106227, 2022.

[5] D. Chalmers, N. G. MacKenzie and S. Carter, "Artificial Intelligence and Entrepreneurship: Implications for Venture Creation in the Fourth Industrial Revolution," *Entrepreneurship Theory and Practice,* vol. 45, no. 5, pp. 1028-1053, 2021.

[6] E. Firt, "Calibrating machine behavior: a challenge for AI alignment," *Ethics and Information Technology,* vol. 25, p. 42, 2023.

[7] W. Lamine, A. Fayolle, S. Jack and D. Audretsch, "Impact of digital technologies on entrepreneurship: Taking stock and looking forward," *Technovation,* vol. 126, p. 102823, 2023.