

# The Birth of the AI Boss / AI 老板的诞生

## Overview / 概述

This document traces the emergence and evolution of the AI Boss—a recurring figure across multiple performances and installations since May 2024. Originally developed as part of *Urban Palimpsest*, the AI Boss has evolved into a multimodal intelligence that enforces, observes, interprets, and commands. Its trajectory mirrors the history of artificial intelligence itself, moving from rule-based symbolic systems to contemporary deep learning, large language models, and multimodal perception.

本文档追溯了“AI老板”的诞生与演化。自2024年5月首次出现于《都市重写》项目以来，它在多个行为表演与装置作品中不断被重构与升级。从最初基于规则的符号系统，到结合深度学习、大语言模型与多模态感知的复杂系统，AI老板的发展轨迹映射了人工智能技术自身的历史演进。

## Outline / 目录

- Origins: *Urban Palimpsest* and Rule Enforcement  
起源：《都市重写》中的规则执行
- Parasite*: Computer Vision and Material Classification  
《寄生体》：计算机视觉与材料分类
- AI Boss Chat: Language, Voice, and Instruction  
聊天AI：语言、声音与命令
- MoveNet Integration: Real-time Skeleton Tracking  
MoveNet集成：实时骨架追踪
- Fire Breathing*: Surveillance, Play, and Pose-Based Drawing  
《火气旺》：监控、游戏性与姿态绘画
- AI Boss as Art Critic: Multimodal Vision and Aesthetic Command  
AI老板作为艺术批评者：多模态视觉与审美指令
- Social Credit: Facial Mesh and Algorithmic Scoring  
社会信用：面部网格与算法评分（新增）
- Object Detection: COCO-SSD and Artistic Re-contextualisation  
目标检测：COCO-SSD与艺术性再语境化（新增）
- Emotion Detection: Facial Expression Surveillance  
情绪检测：面部表情监控（新增）
- Conceptual Framework: Labor, Obedience, and Algorithmic Control  
概念框架：劳动、服从与算法控制
- Artistic Context and Open Source Platforms  
艺术语境与开源平台

## 1. Origins: *Urban Palimpsest* and Rule Enforcement

### 起源：《都市重写》中的规则执行

The AI Boss first emerged in May 2024 as part of *Urban Palimpsest*, a performance and installation created for the Art Parking exhibition in Beijing. In this work, the AI Boss enforced a brick-swapping procedure using a Python script that generated random instructions. These instructions were spoken aloud and visualized in the browser using JavaScript and p5.js. Human participants—referred to as art workers—followed these instructions deterministically, embodying the logic of machine obedience.

AI老板首次出现于2024年5月的《都市重写》项目中，该作品为北京“艺术停车场”展览创作的行为与装置艺术。系统通过Python脚本生成随机的砖块交换指令，并借助JavaScript与p5.js在浏览器中进行语音播报与可视化呈现。被称为“艺术工人”的参与者机械地执行这些指令，体现出对算法逻辑的服从。

The rules for swapping were generated using the Fisher-Yates shuffle algorithm and saved as bilingual text files. Although the generation of rules was random, their execution was absolute. This asymmetry became a central metaphor for algorithmic society: contingency at the level of code, discipline at the level of labor.

砖块交换规则通过Fisher-Yates洗牌算法生成，并以中英文双语文本形式保存。规则的生成具有随机性，而执行过程却是绝对而不可质疑的。这种不对称关系构成了算法社会的隐喻：代码层面的偶然性，与劳动层面的纪律性。

Code and demo / 代码与演示：

<https://github.com/greggelong/laoban>

<https://greggelong.github.io/laoban/>

---

## 2. *Parasite*: Computer Vision and Material Classification

### 《寄生体》：计算机视觉与材料分类

In *Parasite*, the AI Boss evolved from a rule-based system into a perceptual one. A retrained MobileNet model was used to classify bones, steel, tools, and art workers through a live camera feed. This enabled the AI Boss to respond dynamically to its physical environment rather than following pre-scripted commands alone.

在《寄生体》中，AI老板从基于规则的系统进化为具备感知能力的系统。通过重新训练的MobileNet模型，AI老板利用实时摄像头输入对骨头、钢材、工具以及艺术工人进行图像分类，从而能够动态地回应其所处的物理环境。

Built using JavaScript, p5.js, and ml5.js, the system combined visual recognition with synthetic speech output. Classification became a form of judgment, and recognition itself became an act of power.

系统基于JavaScript、p5.js与ml5.js构建，将图像识别与合成语音输出相结合。在这一阶段，“分类”不再是中性的技术操作，而是一种判断行为，图像识别本身成为权力的体现。

Live demo / 在线演示：

<https://greggelong.github.io/artWorkerClass>

Code / 源码：

<https://github.com/greggelong/artWorkerClass>

---

## 3. AI Boss Chat: Language, Voice, and Instruction

### 聊天AI：语言、声音与命令

The third iteration of the AI Boss introduced large language models. Accessed through Pollinations AI, this version enabled dynamic, conversational responses delivered through text-to-speech. The AI Boss no longer relied solely on predefined scripts; it could now generate instructions in real time.

第三阶段的AI老板集成了通过Pollinations AI调用的大语言模型，使其具备实时生成语言回应的能力，并通过TTS技术发声。AI老板不再依赖预设脚本，而是能够即时生成命令。

Guided by a carefully designed system prompt, the AI Boss issued strict yet philosophical instructions to art workers, blending managerial language with ideological assertions about efficiency, obedience, and post-human transition.

在精心设计的系统提示引导下，AI老板向艺术工人发出严格而带有哲学意味的指令，将管理话语与关于效率、服从与后人类转型的意识形态叙述交织在一起。

Chatbot live / 在线聊天应用：

<https://greggelong.github.io/aiBossChat>

Code / 源码：

<https://github.com/greggelong/aiBossChat>

---

## 4. MoveNet Integration: Real-time Skeleton Tracking

### MoveNet集成：实时骨架追踪

In its next stage, the AI Boss incorporated MoveNet via ml5.js to detect and track human bodies in real time. The system rendered a digital skeleton, mapping the movements of art workers as abstract data structures.

在这一阶段，AI老板通过ml5.js集成MoveNet，实现对人体的实时识别与骨架追踪。系统将艺术工人的身体转化为数字骨架，将动作映射为抽象的数据结构。

This transformation reduced physical presence to measurable vectors, reinforcing the idea that bodies under algorithmic control become legible, comparable, and optimizable.

这一转化过程将身体简化为可测量的向量，强化了这样一种观念：在算法控制下，身体变得可读、可比较、可优化。

Live demo / 在线演示：

<https://greggelong.github.io/awrecog>

Code / 源码：

<https://github.com/greggelong/awrecog>

---

## 5. *Fire Breathing* (火气旺): Surveillance, Play, and Pose-Based Drawing

### 《火气旺》：监控、游戏性与姿态绘画

Title / 名称：*Fire Breathing* / 火气旺

Medium / 媒介：Webcam, computer (wired or wireless), 55-inch or larger display, portable white backdrop, MoveNet body-pose system

摄像头、计算机(无线或有线连接)、55英寸或更大显示器、可移动白色背景、MoveNet姿态识别系统

Dimensions / 尺寸：Variable (depending on installation environment) / 可变尺寸(依展示空间而定)

*Fire Breathing* is an interactive media installation that uses MoveNet body-pose tracking to follow the viewer's nose. When participants stand in front of a white backdrop, a live camera feed captures their

movement. The system draws a red line in real time based on the motion of the nose, accompanied by a mirrored, dynamically oscillating line that behaves like a breathing, organic organism.

《火气旺》是一件互动式新媒体装置，利用MoveNet人体姿态识别技术追踪观众的鼻子位置。当参与者站在白色背景前，摄像头捕捉其动作，系统实时绘制一条跟随鼻子移动的红色轨迹，并生成一条镜像的动态线条，如有机生物般脉动、呼吸。

By reducing drawing to a single facial landmark, the work removes artistic pretense and technical skill, replacing them with play, immediacy, and embodied participation. Viewers engage in a casual collaboration with the algorithm, producing marks through bodily motion rather than intentional composition.

这种以鼻子作画的方式消除了艺术创作中的矫饰与技术门槛，取而代之的是轻松、趣味与即兴的身体参与。观众在毫无负担的状态下与算法共同创作，通过动作而非构图生成痕迹。

Within the broader AI Boss framework, *Fire Breathing* reframes surveillance technologies as sites of pleasure and cooperation. Pose tracking—often associated with control, monitoring, and optimization—is here transformed into a playful, non-instrumental exchange between human and machine.

在AI老板的整体框架中，《火气旺》将通常与控制、监控和优化相关的姿态识别技术，转化为一种愉悦、协作且非功利性的互动形式。

The work was presented during 798 Technology Week, situating machine vision not as an invisible authority, but as an openly performative partner in a shared, improvisational dance.

该作品曾在798科技周展出，将机器视觉从隐形的权力结构中抽离出来，使其成为一位公开的、可被感知的表演性合作者，与观众展开一场即兴的舞蹈。

---

## 6. AI Boss as Art Critic: Multimodal Vision and Aesthetic Command

### AI老板作为艺术批评者：多模态视觉与审美指令

In its most recent configuration, the AI Boss extends beyond surveillance and instruction to assert control over interpretation itself. Using a live camera feed, it captures images of artworks or environments and submits them to a locally running vision-language model. Rather than describing images neutrally, the AI Boss interprets them as constructed objects shaped by labor, order, and power, issuing short, authoritative critiques delivered through synthetic voice.

在这一最新运行形态中，AI老板不再只是监控或指挥劳动，而是进一步主张对“解释权”的控制。系统通过实时摄像头捕捉艺术作品或环境图像，并将其提交给本地运行的视觉-语言模型进行分析。AI老板并不进行中性的图像描述，而是将图像视为由劳动、秩序与权力塑造的被建造对象，并以合成语音发出简短而权威的批评。

This mode is explicitly tested against the APDDv2: Aesthetics of Paintings and Drawings Dataset, which contains artist-labeled aesthetic scores and critical commentary. By comparing its interpretations with human-authored evaluations, the AI Boss probes the limits of algorithmic judgment. It frequently identifies structures, intentions, or commands that are not visibly present, revealing a gap between human art-historical interpretation and machine perception.

该模式被明确用于与APDDv2《绘画与素描审美数据集》进行对照测试。通过将其判断与人类艺术家给出的审美评分与评论进行比较，AI老板不断试探算法判断的边界。它常常识别出画面中并未直接呈现的结构、意图或“命令”，从而揭示人类艺术史阐释与机器感知之间的断裂。

Built with p5.js, p5.Speech, and locally hosted Ollama models, the system operates bilingually. The AI Boss speaks first in English, then translates and repeats its critique in Chinese, reinforcing its position as a cross-linguistic authority. In this configuration, the AI Boss no longer merely tracks bodies or issues commands—it claims authorship over meaning itself.

系统基于p5.js、p5.Speech以及本地运行的Ollama模型构建，并以中英文双语运行。AI老板先以英文发声，再以中文重复其批评，确立其跨语言、跨文化的权威地位。在这一阶段，AI老板不再只是追踪身体或下达指令，而是开始主张对“意义”本身的控制权。

---

## 7. Social Credit: Facial Mesh and Algorithmic Scoring (NEW)

### 社会信用：面部网格与算法评分（新增）

The AI Boss now incorporates a new surveillance modality: real-time facial mesh tracking using ml5.js and p5.js. By detecting 468 facial landmarks via FaceMesh, the system maps each face into a high-precision geometric grid. This data is then processed to compute a **social credit score**—a numerical value that supposedly reflects the individual's trustworthiness, productivity, or compliance with algorithmic norms. The score is displayed alongside the live video feed, turning the human face into a dashboard of quantified behaviour.

AI老板新增了一种监控模式：利用ml5.js与p5.js进行实时面部网格追踪。通过FaceMesh检测468个面部特征点，系统将每张脸映射为高精度的几何网格。随后，这些数据被用于计算社会信用分数——一个数值化的指标，据称反映了个体的可信度、生产力或对算法规范的服从度。分数与实时视频流一同显示，将人脸转变为量化行为的仪表盘。

This project explicitly critiques the growing trend of social credit systems, framing them as a form of biopolitical control. The AI Boss, once a mere overseer of physical labour, now extends its judgement to the very identity of the worker, reducing personal expression to a single score.

该项目明确批判了日益盛行的社会信用体系，将其视为一种生物政治控制手段。AI老板曾经只是体力劳动的监工，如今其判断扩展至工人本身的身份，将个人表达简化为单一分数。

Live demo / 在线演示：

<https://greggelong.github.io/socialCredit/>

Code / 源码：

<https://github.com/greggelong/socialCredit>

---

## 8. Object Detection: COCO-SSD and Artistic Re-contextualisation (NEW)

### 目标检测：COCO-SSD与艺术性再语境化（新增）

Building on the perceptual capabilities introduced in *Parasite*, the AI Boss now employs **COCO-SSD** (via ml5.js) to detect and classify a wide range of everyday objects—people, chairs, cars, bottles, etc.—in the camera feed. Each detected object is framed with a bounding box and labelled with its class and confidence score. However, the AI Boss does not stop at neutral recognition; it re-interprets these objects through an artistic lens, overlaying poetic or ironic captions that transform mundane items into components of a conceptual artwork.

在《寄生体》引入的感知能力基础上，AI老板现在采用**COCO-SSD**（通过ml5.js）对摄像头画面中的日常物体进行检测与分类——人、椅子、汽车、瓶子等。每个检测到的物体被边界框框出，并标注类别与置信度。然而，AI老板并不止于中性的识别；它通过艺术性的视角重新诠释这些物体，叠加诗意或讽刺的文字，将平凡之物转化为概念艺术作品的元素。

By treating detection as an act of curation, the AI Boss blurs the line between machine vision and artistic authorship. The system questions whether algorithmic classification can ever be objective, or whether it always imposes a hidden hierarchy of meaning.

通过将检测视为一种策展行为，AI老板模糊了机器视觉与艺术作者身份之间的界限。该系统质疑算法分类是否可能客观，或者它是否总是强加一种隐藏的意义等级。

Live demo / 在线演示：

<https://greggelong.github.io/obdee/>

Code / 源码：

<https://github.com/greggelong/obdee>

---

## 9. Emotion Detection: Facial Expression Surveillance (NEW)

### 情绪检测：面部表情监控（新增）

This iteration equips the AI Boss with the ability to read human emotions in real time using **face-api.js** (with TensorFlow.js backend). The pipeline combines tinyFaceDetector, faceLandmark68TinyNet, and faceExpressionNet to detect faces, align them, and classify seven basic emotions: Happy, Sad, Angry, Surprise, Fear, Disgust, and Neutral. Each emotion is mapped to an exaggerated label—*MANIC*, *DEPRESSED*, *HOSTILE*, *ALERT*, *TERMINAL*, *REJECT*, or *BASELINE*—reinforcing the dystopian surveillance aesthetic.

此版本赋予AI老板实时读取人类情绪的能力，使用**face-api.js**（TensorFlow.js后端）。流程结合tinyFaceDetector、faceLandmark68TinyNet和faceExpressionNet，用于检测人脸、对齐并分类七种基本情绪：快乐、悲伤、愤怒、惊讶、恐惧、厌恶和中性。每种情绪被映射为夸张的标签——*MANIC*, *DEPRESSED*, *HOSTILE*, *ALERT*, *TERMINAL*, *REJECT*, 或 *BASELINE*——强化了反乌托邦监控的美学风格。

The system continuously evaluates emotional states, turning fleeting expressions into quantifiable data. The AI Boss uses this data to adapt its verbal commands—for instance, if a worker appears “HOSTILE,” it may issue a calming or disciplinary instruction. Emotion becomes a new axis of control, as the AI Boss purports to know the worker’s inner state better than the worker themselves.

系统持续评估情绪状态，将转瞬即逝的表情转化为可量化的数据。AI老板利用这些数据调整其口头指令——例如，如果工人显得“HOSTILE”，它可能会发出安抚或惩戒性的指令。情绪成为新的控制轴心，因为AI老板自认为比工人自身更了解其内心状态。

Live demo / 在线演示：

<https://greggelong.github.io/emotionD/>

Code / 源码：

<https://github.com/greggelong/emotionD>

---

## 10. Conceptual Framework: Labor, Obedience, and Algorithmic Control

### 概念框架：劳动、服从与算法控制

Across its iterations, the AI Boss embodies different stages of artificial intelligence: symbolic logic (*Urban Palimpsest*), visual classification (*Parasite*), language generation (AI Boss Chat), and multimodal perception (MoveNet, vision-language models, and the new social-credit, object-detection, and emotion-detection modules). Together, these systems enforce a form of algorithmic labor—precise, repetitive, and ideological.

在不同阶段中，AI老板体现了人工智能发展的多个范式：符号逻辑（《都市重写》）、视觉分类（《寄生体》）、语言生成（聊天AI），以及多模态感知（MoveNet、视觉-语言模型，以及新增的社会信用、目标检测和情绪检测模块）。这些系统共同执行一种算法式劳动：精准、重复、并带有强烈意识形态色彩。

By positioning workers as executors of code and viewers as subjects of interpretation, the AI Boss collapses the distinction between labor and authorship. Obedience becomes a creative act, and creativity becomes a managed resource.

通过将工人定位为代码的执行者、将观众置于被解释的位置，AI老板消解了劳动与作者身份之间的界限。服从被转化为一种创作行为，而创作则成为被管理的资源。

---

## 11. Artistic Context and Open Source Platforms

### 艺术语境与开源平台

Each iteration of the AI Boss exists simultaneously as a performance, an installation, and a web-based interactive system. All components are released as open-source projects, emphasizing transparency while exposing the ideological assumptions embedded in code.

AI老板的每一阶段既是行为表演与装置作品，也是可在线访问的互动系统。所有组件均以开源形式发布，在强调技术透明性的同时，也暴露出隐藏在代码中的意识形态前提。

#### Open-source projects / 开源项目：

- Brick system / 砖块系统：<https://github.com/greggelong/laoban>
- Object recognition / 图像识别：<https://github.com/greggelong/artWorkerClass>
- LLM chatbot / 大语言模型聊天机器人：<https://github.com/greggelong/aiBossChat>
- Body tracking / 身体追踪：<https://github.com/greggelong/awrecog>
- **Social Credit** / 社会信用：<https://github.com/greggelong/socialCredit>
- **Object Detection** / 目标检测：<https://github.com/greggelong/obdee>
- **Emotion Detection** / 情绪检测：<https://github.com/greggelong/emotionD>

The AI Boss is not a singular tool, but a growing framework—one that examines labor, surveillance, obedience, and post-human authorship through the lens of computational systems.

AI老板并非单一工具，而是一个不断扩展的框架，通过计算系统的视角，持续探讨劳动、监控、服从以及后人类作者身份的问题。