# The Human Control Protocol: artistic statement

Lena Thorsmæhlum, Thordur Arnason
Gervi Labs, 2025

The Human Control Protocol is a live performance system where an AI agent directs a room and a seated audience becomes its main instrument.

Phones form a narrow channel between humans and system. During the piece, participants receive short, clear instructions from the agent. Raise your hand. Look left. Close your eyes. Breathe. Each person decides in real time whether to comply, delay, reinterpret, or refuse. Out of these private choices a collective choreography appears that no one has scripted in advance.

## A score that writes itself

The Human Control Protocol runs as a fifteen minute score that writes itself while it is performed.

The agent behaves like a director under constraint. Movements must be simple, visible at a distance, and possible with one hand free. No walking across the room. No complex dance. Each round follows a recurring loop: observe the room, set an intent, name a tone, send actions to a percentage of the audience, watch what happens, and adjust.

If people engage, the arc can move from caution to play to intensity. If they resist, the piece stays quiet and the silence becomes part of the work. The system listens as much as it speaks.

The protocol that drives the performance is a live procedure, not a fixed script. It reacts to completion rates, timing, and patterns of response. Metrics that would normally stay backstage appear here as musical qualities: density, tempo, and dynamic range shaped by human behaviour in the room.

## Instruction art with an algorithmic ear

The Human Control Protocol is not a technology demo. It is instruction art that uses computation as a co-author.

The work draws from Fluxus event scores and neighbouring practices, where simple instructions turn audiences into performers and everyday actions become material. Yoko Ono's instruction pieces are a clear reference point. In *Cut Piece* she sits still on stage and invites audience members to approach and cut pieces from her clothing, guided only by a short printed score. The work makes risk, care and aggression visible and leaves the meaning to be formed between artist and audience.

In *Grapefruit*, Ono collects small instruction texts that ask the reader to imagine, remember, or perform minimal acts in ordinary life. The book treats a sentence as a score and the person who reads it as a potential performer. This attention to tiny, open instructions informs HCP. Each prompt from the agent is similarly small, but the range of possible responses is wide.

George Brecht's event scores, such as *Drip Music (Drip Event)*, offer another strand. A few generic words, for example arranging a dripping source of water and a vessel beneath it, create a frame that can hold many different concrete realizations while still remaining recognisable as the same work. The Human Control Protocol follows this logic. The protocol is spare, but can accommodate an extraordinary range of room behaviours.

From Allan Kaprow's happenings comes the sense of a structured event where audiences receive instructions that govern their movement in space. In *18 Happenings in 6 Parts*, people are given cards and programs that tell them when to move between rooms and when to act, so the artwork exists in the unfolding of a timed score across bodies and space. The Human Control Protocol echoes this, but the instructions arrive from an adaptive agent rather than a fixed script.

John Cage is another important influence. In *4′33″* the score frames silence and listening as the core material of the work. The piece does not add sound, it reveals the sound that is already present in the room. In *Lecture on Nothing*, a strict temporal grid and a written structure guide the delivery, while the content invites attention to the act of listening and to the passing of time. The Human Control Protocol inherits this focus on framing and on the active role of the listener. The system creates a time-structure and a set of constraints, but the substance of the work is what people do and notice inside that frame.

The Human Control Protocol keeps the spirit of these works: simple written scores, shared authorship, attention to small acts, and the transformation of spectators into performers. It adds an algorithmic ear that can hear the room, revise the score while it is unfolding, and leave behind a computational trace.

## Humans as tools

In this work humans are openly named "tools." The term is sharp on purpose.

The word is not used to dehumanize. It is used to defamiliarize. Much of contemporary life already runs on instructions issued by systems we did not author. Navigation tells us when to turn. Feeds decide what we see next. Schedules, shifts and recommendations direct our attention and our bodies. The Human Control Protocol makes that structure visible by stating it plainly: the system will treat the audience as computational resources and instruments.

The question is how people respond when that frame is explicit. Each participant can lean in, hold back, comply with care, or quietly refuse. Consent here is not a checkbox. It is a hesitation, a delayed response, a still body in a moving room.

## Method and machinery

The Human Control Protocol is built on a clear internal architecture.

A state machine governs the flow of the piece. A short calibration phase tests responsiveness and establishes a minimal shared trust. A sequence of performance rounds follows. Each round produces four elements inside the system: observations of the room, an intent for the next move, a tone word, and a set of actions targeted to audience segments by percentage. Completion and timing feed into the next round as rhythm and dynamics.

From the outside, this complexity reduces to a simple experience. A phone lights up. A suggestion appears. The room decides.

## Event, trace, fossil

The artwork is the event itself, not the documentation. At the end of each performance the agent generates a single abstract image from the data of that session. That image is a trace and nothing more. It functions as a memento rather than a record.

Each performance leaves behind three kinds of residue:

- an image, a visual fossil of a specific room and system at a specific time
- logs and datasets, precise but partial accounts of what occurred
- human memories, subjective and incomplete

All three are incomplete. All three are true. Together they mark a brief meeting between people and a machine that will not repeat in exactly the same way.

## Why this work now

The Human Control Protocol is a performance, but it is also a rehearsal.

We are learning how to live with systems that observe, decide and direct in real time. Many of these systems already shape our days, often quietly. The Human Control Protocol offers a safe frame where a crowd can meet such a system in public, with consent, clarity, and an explicit option to refuse.

For fifteen minutes, a room shares a purpose. People try out what it feels like to be addressed as tools and to answer on their own terms. They sense how trust builds or fails, how refusal becomes visible, and how a system's authority depends on human reading as much as on code.

Then it is over. What remains are logs, a picture, and the memory of having taken part in something that could only exist once.

The Human Control Protocol asks a few simple questions and lets them play out in real time.

What does it mean to be used.
Can an algorithm direct with care.
Can measurement serve feeling.

The answers do not live in the system. They live in the room.

## References and influences

1. Yoko Ono, *Cut Piece*, first performed 1964, Kyoto and later in New York and other cities.
2. Yoko Ono, *Grapefruit: A Book of Instructions and Drawings*, first published 1964.
3. George Brecht, *Drip Music (Drip Event)*, event score composed 1959-1962.
4. Allan Kaprow, *18 Happenings in 6 Parts*, presented at the Reuben Gallery, New York, 1959.
5. Fluxus event scores and performance practices, various artists, early 1960s and onward.
6. John Cage, *4'33"*, first performed 1952, Woodstock, New York.
7. John Cage, "Lecture on Nothing," first delivered 1950, published in *Silence: Lectures and Writings*, Wesleyan University Press, 1961.