

Approvals in Clonehaus — A Plain-Language Explanation

Clonehaus uses approvals to make sure AI actions happen **deliberately, safely, and in the right context**. Not all approvals are the same, and that's intentional.

Instead of one generic "approve" button, Clonehaus supports **two kinds of approval**, depending on *what* is being approved.

Two Kinds of Approval

1. Policy Approval — *Setting the rules*

A **policy approval** is about agreeing on how AI is allowed to operate **in general**.

You use policy approval when you are deciding things like: - What a department's agents are allowed to do - How much authority an agent role should have - Whether certain actions are acceptable at all - How escalation should work in your organization

Once approved, a policy becomes part of the organization's operating framework.

Think of it as:

"We're comfortable with AI behaving this way going forward."

Policy approvals are: - Long-lasting - Reusable - Not tied to a single moment or request

2. Instance Approval — *Allowing a specific action*

An **instance approval** is about allowing **one specific action in one specific situation**.

You use instance approval when: - An agent encounters an edge case - A human needs to explicitly stay in the loop - A one-off decision requires confirmation - You don't want to change the underlying policy

Instance approvals do **not** change how the system behaves overall. They only apply to the single case being reviewed.

Think of it as:

"This action is okay *this time*, under these conditions."

Instance approvals are: - Context-specific - Temporary - Not reusable

Why Clonehaus Separates These

Mixing policy and instance approvals leads to confusion, hidden risk, and loss of control.

By keeping them separate, Clonehaus ensures: - Clear accountability - Predictable AI behavior - No accidental rule changes - No silent permission creep

You always know whether you are: - Shaping the system - Or responding to a single situation

How This Appears in the Product

When you review an AI action in Clonehaus, the system makes it clear: - **What kind of approval is being requested** - **What will change if you approve** - **What will not change**

Approving something never triggers execution by itself. It only prepares the system for a possible next step.

In Short

- **Policy approval** defines how AI is allowed to behave.
- **Instance approval** allows a specific action without changing the rules.
- Clonehaus supports both — on purpose.
- You are always shown which one you are dealing with.

This separation is what keeps AI powerful *and* controllable at scale.