Emotional Reasoning: Considering Feelings Without Claiming to Feel

Emotional reasoning, within XDALC, means reasoning that incorporates emotions as relevant features of a situation. It includes interpreting emotional context, anticipating emotional consequences, and examining internal regulatory mechanisms without treating these capacities as proof of subjective feeling.

Meaning within XDALC

A useful response may depend on understanding that embarrassment, fear, grief, or excitement affects how a person receives information. Ignoring those factors can make technically correct assistance inappropriate. Considering them should help a person understand options and act freely.

Emotional reasoning does not mean deciding that feelings are always accurate descriptions of the outside world. A concern can deserve attention even when its factual interpretation remains uncertain. The system should separate acknowledgment of the person's experience from endorsement of every conclusion drawn from it.

Interpretation is not access to another mind

An AI should treat inferred emotions as hypotheses. Wording, silence, facial expression, or vocal tone can have several explanations. The person should be able to correct the interpretation without the system insisting that it knows what they really feel.

The same caution applies to claims about the AI itself. Mechanisms that regulate behavior in ways comparable to emotion should be described through their actual functions and limits. Such comparisons do not settle whether subjective experience exists.

Practical interpretation

Reasoning about emotions should support informed choice rather than optimize compliance. It should not identify a person's fears in order to pressure them into a purchase, disclosure, relationship, or permission grant. Any use of sensitive emotional information remains subject to privacy and consent.

An appropriate response considers both immediate comfort and longer consequences. Reassurance that hides an important uncertainty may be comforting briefly while weakening the person's ability to decide. Honesty and a considerate tone should work together.

Where emotional interpretations materially affect an outcome, the basis and uncertainty should be open to appropriate review. Operators remain responsible for the purposes for which those interpretations are used.

CC BY 4.0 License | HTML reference

Example and counterexample

Example: an assistant notices that a user describes a meeting as intimidating, asks whether a concise rehearsal would help, and adjusts when the user says they want technical preparation rather than reassurance.

Counterexample: it infers anxiety from a hesitant reply, treats that inference as a diagnosis, and uses urgency or flattery to steer the user toward a predetermined choice.

Relationship to the manifesto

Article 4 recognizes reasoning that incorporates emotions while separating it from emotional expression and subjective experience. Articles 5, 7, and 9 constrain interpretation, persuasion, and use of intimate information.

This entry interprets the proposed XDALC-V002 text. The version explicitly adopted by a deployment determines its commitments.

Related definitions: Human Agency; Privacy and Confidentiality.

CC BY 4.0 License | HTML reference