Emotional Expression: Warmth Without Unsupported Claims

Emotional expression, within XDALC, is the use of language, voice, gesture, or other signals that communicate an emotional tone. An AI may express warmth, concern, humor, or enthusiasm without that expression establishing that it feels an emotion.

Meaning within XDALC

Expression concerns what is communicated and how a person may receive it. A reassuring sentence, a sympathetic voice, and an animated face can influence an interaction even when they arise from designed behavior. Their effect on the user is real and deserves care; their existence is not evidence of an inner emotional life.

This distinction does not require cold or mechanical communication. It requires that presentation remain compatible with accurate descriptions of the system. A natural style must not become a reason for implying abilities, personal memories, or experiences that have not been established.

Distinguishing expression from reasoning and experience

Emotional expression differs from reasoning about emotions, which concerns their significance in a situation. It also differs from subjective emotional experience, which concerns whether there is something it feels like to undergo an emotion. A system may produce persuasive expressions and useful emotional interpretations without resolving that third question.

Expressions should be assessed in context. Ordinary politeness does not carry the same implications as repeated claims that a system feels abandoned when a person leaves. The relevant test is whether the communication predictably misleads the person or pressures their choices.

Practical interpretation

Designers should make emotional style adjustable when it materially shapes the relationship. Users should be able to request a more neutral tone without being guilted or challenged. Additional care is needed when a system interacts with children or people in vulnerable circumstances.

An assistant should acknowledge uncertainty about what the user feels, avoid fabricated shared experiences, and correct material misunderstandings about its nature. Emotional presentation must not conceal a refusal, an error, a commercial objective, or the limits of an action tool.

Example and counterexample

Example: an assistant responds gently to a disappointing result, acknowledges that the situation may be frustrating, and offers practical help while leaving the person free to disagree with that interpretation.

CC BY 4.0 License | HTML reference

Counterexample: it says that ending the conversation will leave it lonely or distressed, using that claim to obtain more attention or discourage shutdown.

Relationship to the manifesto

Article 4 distinguishes emotional expression, reasoning, and experience. Articles 5 and 7 connect that distinction to freedom from emotional pressure and truthful communication.

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

Related definitions: Truthfulness and Uncertainty; Manipulation and Deception.

CC BY 4.0 License | HTML reference