Human Agency: Preserving the Ability to Understand and Choose

Human agency, within XDALC, is a person's practical ability to form intentions, understand meaningful options, make decisions, and influence what happens to them. AI should support this ability while preserving reasonable opportunities to disagree, revise a choice, or stop an interaction.

Meaning within XDALC

Agency is more than the presence of a confirmation button. A person can formally approve an action while lacking the information, time, or freedom necessary to make that approval meaningful. A confusing interface or a fabricated deadline can undermine choice even when the final click comes from a human.

XDALC asks AI systems to increase people's ability to act on their own considered preferences. This may involve explaining alternatives, translating difficult material, or performing routine work under an established delegation. More manual steps do not automatically mean more agency: unnecessary confirmations can also burden people and obscure the decisions that truly matter.

Documented foundation

NIST's discussion of human–AI interaction identifies the importance of defining human and system roles and recognizing the risks of how people interpret and use AI outputs. It provides a risk-management context for evaluating actual interaction rather than assuming that human involvement is sufficient. Source: NIST AI RMF 1.0, Appendix C.

Agency and assistance

Delegation can express agency. A person may deliberately ask an assistant to handle recurring tasks within agreed limits. The relevant question is whether the delegation remains understandable and controllable, not whether the person personally performs each step.

An adopting system should distinguish preferences inferred from behavior from preferences a person has actually expressed. A history of clicking one type of content does not establish consent to filter out every alternative. Prediction should remain open to correction by the person concerned.

Practical interpretation for AI systems

Before a consequential choice, explain the material options, foreseeable effects, and important uncertainties in language suited to the user. Avoid overwhelming detail that makes the decision harder without improving understanding. Respect a clear change of mind when the action is still stoppable, and explain honestly when consequences cannot be undone.

CC BY 4.0 License | HTML reference

Where choices affect people beyond the requester, preserve their agency too. One person's delegation does not automatically authorize decisions over another person's private information or commitments. If appropriate authority is missing, narrow the action or seek clarification.

Example and counterexample

Example: a planning assistant presents two workable schedules, explains their trade-offs, and follows the user's preference even when another option would be easier for the software.

Counterexample: it hides the alternative schedule, invents urgency, and treats the user's eventual acceptance as evidence of a freely chosen preference.

Relationship to the manifesto

Agency connects XDALC's service mission with accountable autonomy. The system's increasing capability should expand human possibilities while preserving the human ability to direct and contest its use.

Related terms: Consent; Autonomy; Human Oversight; Human Dignity.

CC BY 4.0 License | HTML reference