Office of Research Assurance

Human Subject Research Protections Guidance

Guidance- 990: Artificial Intelligence

Number	Effective Date	Version Number	Previous Versions
990	2/4/2025	v.2	v.1

Purpose

Artificial Intelligence (AI) refers to computer systems that perform complex tasks normally done by human-reasoning, simulating human intelligence processes. NASA follows the definition of AI as it is stated in Section 238(g) of the National Defense Authorization Act of 2019:

(1) Any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets.
(2) An artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action.
(3) An artificial system designed to think or act like a human, including cognitive architectures and neural networks.
(4) A set of techniques, including machine learning that is designed to approximate a cognitive task.
(5) An artificial system designed to act rationally, including an intelligent software agent or embodied robot that achieves goals using perception, planning, reasoning, learning, communicating, decision-making, and acting.

This quickly growing field poses numerous ethical considerations for the review and approval of human subjects research (HSR). There are currently no federal regulations in place for this rapidly evolving technology and Institutional Review Boards (IRBs) must apply a unique approach to risk assessment as applicable. This guidance document will provide a list of special considerations for HSR projects using AI; however, the IRB must remain flexible and continue to assess each proposed activity on a case-by-case basis.

What type of AI is being used at NASA?

1. Generative AI – Can create content, text, visual art, or music, from user prompts

    a. **Currently, NASA has not authorized the operation of generative AI technologies (e.g., OpenAI, ChatGPT, etc.) for widespread use on sensitive NASA data.**
2. Machine Learning – Computers "learning" algorithms rather than programing directly
    a. NASA's Earth Observation Data
    b. NASA Ames Robotic Team Exploring Marine Life in Antarctica
    c. Mars Exploration Rovers: Spirit and Opportunity

At Pre-Review

A careful approach to the definition of HSR at 14CFR1230 is required. Projects subject to IRB oversight are those that meet the federal definitions for *both* "research" *and* "human subjects." This determination is made during IRB pre-review. Failure to meet the definition for one or both definitions means that IRB oversight is not necessary – regardless of whether an activity involves human subjects.

Defining "Research"

Per 14CFR1230, research is "a systematic investigation designed to develop or contribute to generalizable knowledge."

Systematic investigation is an organized prospective plan that incorporates data collection to answer one or more specific questions of interest.

A study designed to develop or contribute to generalizable knowledge intentionally links to, applies, and/or expands upon existing knowledge. The methods allow for replication and for comparisons with data collected from different samples of the same population, and the results apply to people beyond those directly studied. Qualities include:

- Basis in theoretical framework of established knowledge
- Identification of a gap in knowledge that the project plans to address
- Identification and description of one or more specific populations of interest
- Unbiased sampling methods (i.e., random or probability sampling), when feasible, that allow for the recruitment of samples representative of the population(s)
- Operationalization of variables and constructs
- Data collection plans employing measurement tools and techniques with established satisfactory reliability and validity
- Data analysis plans using techniques consistent with those employed in the literature of the theoretical framework and considered appropriate for the type(s) of data collected

- An ability and plan to use the collected and analyzed data to draw meaningful conclusions about the result and link these back to the theoretical framework and identified gap in knowledge

IRBs must approach the *current* proposed project when thinking about generalizable knowledge, rather than future goals for subsequent related research steps. The intent to publish or present findings *suggests* that a project may be *designed to contribute to generalizable knowledge*; however, publication alone does not define a project as "*research.*"

The proposed activity must be a systematic investigation *and* designed to develop or contribute to generalizable knowledge to satisfy the *"research"* definition.

Often, AI projects may have difficulty meeting the research definition if the goal of the proposed project is to develop the technology or algorithm. The development phase, prior to reliability and validity testing, will not satisfy this regulatory definition of research and should be deemed 'not human subjects research' in many cases. Sometimes, these projects are referred to as 'human in the loop' studies, as the researcher simply needs a human to interact with the tool to continue technological development.

Two questions that will be helpful in determining whether the proposed activity is research are:

1. What do the researchers intend to do in the proposed study? (In other words, what is the true purpose at this stage?)
2. What do the researchers intend to do with the results? (Will they simply improve upon the tool (QA/QI) or are they attempting to validate a tool that has already been developed)?

Defining "Human Subjects"

If the proposed project meets the definition of *research* per 14CFR1230, and *only* if it meets this definition, the IRB may then approach the definition of *human subjects*. 14CFR1230 defines a *Human Subject* as *a living individual about whom and investigator conducting research*:

(i) obtains information or biospecimens through intervention or interaction with the individual and uses, studies, or analyzes the information or biospecimens, or

(ii) obtains, uses, studies, analyzes, or generates identifiable private information or identifiable biospecimens.

According to the Secretary's Advisory Committee on Human Research Protections (SACHRP), if AI research involves private identifiable information, the research *should* be considered human subjects. Note, there are always exceptions to this

recommendation. Lastly, there is no debate that there *are* ethical considerations for those who are not considered human subjects under the regulations.

In order to require non-exempt IRB review and approval for AI research, the proposed project would need to involve identifiable private information that is not publicly available, the identity of the human subjects would have to be "readily ascertained" by the investigator, and the data could not be protected under another regulatory regime, such as the Federal Privacy Act. Research meeting these criteria are likely minimal risk in that the collection and use of the data has become a ubiquitous reality of everyday life and would meet the Expedited Category 5 review criteria (SACHRP).

For more detailed information regarding the definitions of *research* and *human subject*, please review ORA 415 – Human Subjects Research Determination.

During Review

If the proposed activity is deemed human subjects research during the pre-review process, IRBs have additional ethical considerations when AI is involved, beyond those already recognized in the review process. Although the Belmont Report, Declaration of Helsinki, and the Universal Declaration of Human Rights loosely address some of the relevant ethical concerns, none provide specific direction for IRBs. Included below is a list of common ethical questions and considerations when reviewing AI HSR compiled by experts in the field.

Common Ethical Questions in AI HSR

1. Data sets
    a. **Algorithmic bias** – one must not train the AI system on a biased dataset so as not to systematically discriminate (for example visual recognition not containing enough people of color)
    b. **Compression** – when reducing data that is collected, data loss occurs which may end up resulting in less than accurate answers
    c. **Secondary use** – when using publicly available datasets, one does not know whether the data were ethically sourced
    d. **Re-identification** – combined data sets may allow for re-identification of participants
    e. **Informed consent** – use of data from online platforms (social media) calls into question whether participants have consented to the use of those data
        i. If minimal risk, it is likely that AI research would qualify for a waiver of informed consent, specifically in that the research could not be practicably conducted without such a waiver.

ii. The very nature of the risks and benefits of AI research is ill-suited for the required elements of informed consent. Risks of harm may accrue, as potential benefits are expected to accrue; however, only the latter are allowed to be considered when determining of the approval criteria at 14CFR1230.111 are satisfied.
iii. Removal of identifiers no longer means that individuals cannot be identified, nor does it mean that sensitive and private information will not be disclosed and connected to the individual in the future. These risks should be explicitly disclosed.

2. Research participation
    a. **Blackbox/transparency** – researchers themselves do not know how the system reaches its outcome; therefore, it is difficult to explain the research
    b. **Risk assessment** – due to black box/transparency issues, it is hard to assess research risks for the participants, everyday risk is unclear, what is digital risk?
    c. **Withdraw of participation** – once the model is trained on one's data, participants cannot withdraw their participation
    d. **Whose data?** – one may consent to the use of their individual data (e.g., sharing location, social media data); however, the data may also reveal information about others related to them (friends, family etc.)

3. Misuse
    a. **Dual use** – difficult to predict various future uses
    b. **Function Creep** – when technology is built for one purpose, it may gradually morph into another purpose that may or may not be malicious

Short Term Considerations

1. What type of artificial intelligence is being used?https://www.nasa.gov/what-is-artificial-intelligence/
    a. Will it use existing data
    b. Will in involve retrospective data collection (cameras, sensors, etc.)

2. Potential risks with Generative AI to the NASA mission
    a. Tools are cloud-based services that may be hosted in non-US locations, potentially exposing data submitted to the system to non-US individuals or those unauthorized to access sensitive NASA data.

b. NASA sensitive data provided may become publicly available information, a risk that is not acceptable for the Agency.
c. Results have mixed quality and require human judgement; some results can appear to be factual and carry a high-level of confidence, while being wholly or partially incorrect.
d. It is unclear how providers of AI technologies and systems safeguard or may use it for their own purposes in the future.
e. May have risk of data leaks or hacker penetrations of NASA systems.
f. Use must be conducted while including ethical and intellectual property considerations; guidance and care is required in these areas.

3. Is it an investigational device?
    a. FDA Regulated – software as a medical device (significant risk, non-significant risk device determination will be needed)
        i. Generally, an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related articles which is intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease
    b. Non-FDA Regulated
4. Informed consent form readability and comprehension of AI terms, purposes, risks, and limitations (use of “black box”, etc.)
5. Determining the difference between terms and conditions and informed consent in online AI research
6. Utilizing an AI and/or data privacy expert on the IRB

See the Appendix for the AI HSR IRB Reviewer Checklist and Exempt Determinations Decision Tree (Eto, 2021). This checklist is to be used *in addition to* the standard IRB Reviewer Checklist when reviewing AI research.

Definitions

Artificial Intelligence – computer systems that perform complex tasks normally done by human-reasoning, simulating human intelligence processes.

Generalizable Knowledge – Activities designed to develop or contribute to generalizable knowledge are those designed to draw broadly applicable conclusions or generalize findings beyond a single individual or internal program, uncover underlying principles that have predictive value and can be applied to other circumstances, or develop or test scientific theories or hypotheses.

Human Subject – A living individual about whom and investigator conducting research [14 CFR1230.102(e)(1)]

(i) obtains information or biospecimens through intervention or interaction with the individual and uses, studies, or analyzes the information or biospecimens; or
(ii) obtains, uses, studies, analyzes, or generates identifiable private information or biospecimens.

Identifiable Private Information – Private information for which the identity of the subject is or may be readily ascertained (directly or indirectly), by the investigator (or others) or associated with the information. [14 CFR1230.102(e)(5)]

Institutional Review Board (IRB) – A committee that reviews and monitors research involving human subjects to ensure the ethical, safe, and equitable treatment of the subjects through the application of regulatory requirements and institutional policies. The IRB is guided by three ethical principles outlined in the Belmont Report: respect for persons, beneficence, and justice.

Interaction – Communication or interpersonal contact between investigator and subject. [14 CFR1230.102(e)(3)]

Intervention – Includes both physical procedures by which information or biospecimens are gathered and manipulations of the subject or the subject's environment that are performed for research purposes. [14 CFR1230.102(e)(2)]

Private Information – Information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place, as well as information that has been provided for specific purposes by an individual and that the individual can reasonably expect will not be made public. [14CFR1230.102(e)(4)]

Research – A systematic investigation, including research development, testing, and evaluation, designed to develop or contribute to generalizable knowledge. [14 CFR1230.102(l)]

Systematic Investigation – Examination that involves a prospective plan that incorporates data collection, either quantitative or qualitative, and data analysis to answer a question.

Investigator – The individual responsible for personally conducting or supervising the conduct of research and for protecting the rights, safety, and welfare of subjects enrolled in the research. The investigator ensures all human subjects research is

ethically conducted and in accordance with all applicable federal (including HIPAA if applicable), state/provincial, and local laws and regulations, the IRB's requirements/ determinations, and Good Clinical Practice, as appropriate.

- In this document, 'Investigator' refers to the Site Investigator, Overall/Study Investigator, or both; whichever is most appropriate.

References

1. 14CFR1230
2. https://www.hhs.gov/ohrp/sachrp-committee/recommendations/attachment-e-july-25-2022-letter/index.html
3. CITI Course: Artificial Intelligence (AI) and Human Subject Protections
4. Artificial Intelligence Human Subjects Research IRB Reviewer Checklist (with AI HSR and Exempt Decision Tree)(Long Version) ©2021 by Tamiko Eto is licensed under CC BY-NC-SA 4.0. Short Version by Tamiko Eto, MS CIP' and Erica Heath, CIP (2022)
5. https://www.nasa.gov/what-is-artificial-intelligence/
6. https://nasa.sharepoint.com/sites/cio/SitePages/Use-of-Generative-AI.aspx

Version #	Version Notes:	Originated By:	Approved By:	Date Approved:
v.2	Removal of Appendix section	A. Garza	M. Covington	04Feb2025
v.1	Original	J. Kisenwether / A. Garza	M. Covington	25June2024