NC State University IRB Guidance

Use of Artificial Intelligence (AI) and Generative AI (GAI) in Human Subjects Research

This guidance is intended to assist researchers at NC State University in the participant-centered design and implementation of human subjects' research, including completion of the IRB application and risks assessment, wherein any artificial intelligence (hereafter, AI) or generative AI (hereafter, GAI) is used. Researchers should utilize this guidance if planning to design and/or implement research with human subjects where AI or GAI is used to interact with participants, used as an intervention with participants, used to manipulate a participant's environment, or used to access, analyze, or generate information about one or more participant(s).

AI can be used in human subjects' research in several ways. It can help you develop a hypothesis or research question, identify possible participant groups and avenues for recruitment, suggest modes of data collection and generate draft instruments, transcribe and analyze video, audio, and still photos, create case studies or scenarios, analyze data, and so much more. The NC State University IRB does not regulate many aspects

of research utilizing AI, but in many instances, the research using AI will require IRB regulation via review and approval before implementation.

The IRB distinguishes between two types of AI and human subjects' research:

- AI focused on analyzing and interpreting existing data.
- GAI creating new content, such as text, images, or code, and, like other AI, is based on learned patterns from data.

The IRB will regulate research (Google document) with human subjects (Google document) using AI or GAI when the researchers design the study in a way that the AI or GAI:

- Interacts with a participant
- Is used as an intervention with a participant including biomedical or behavioral interventions
- Manipulates the environment of a participant
- Analyzes directly identifiable or indirectly identifiable data about a participant
- Generates information about an individual participant in a study

NC State University-Wide Expectations and Standards on the Use of Generative AI

Resources for an individual who is acting on behalf of NC State University in the scope of their official role, to complete human subjects' research.

- NC State University "AI at NC State" hub (opens in a new window)
- NC State University Guiding Principles for AI Tools (Google document)
- NC State University Approved Enterprise AI Tools (opens in a new window)
- NC State University perspective on Gemini Chatbot (opens in a new window)
- NC State University AI Tool IT Purchase Process (opens in a new window)
- NC State University Graduate Handbook Section 3.6.E, AI Use in Theses and Dissertations (opens in a new window)
- NC State University AI Tools for Certain Data Sensitivity Levels (opens in a new window)
    - Note, if you want to use any of the AI Tools noted above for "Red" or "Purple" data, you will need to initiate the IT Purchase Compliance (opens in a new window) process as noted in the link AI Tools for Certain Data Sensitivity Levels.

Relevant IRB Guidance

- Risks and Benefits in Human Subjects Research (Word document)
- Applications and Software in Human Subjects Research (Word document)
- Use of Social Media for Human Subjects Research (Word document)
- Benign Behavioral Interventions (Word document)
- Deception in Human Subjects Research (Word document)
- Secondary Data and the IRB (Word document)
- Feasibility and Pilot Work (Word document)
- Phased and Staged Research (Word document)
- Medical Devices (Word document)

- Clinical Trials (Word document)
- FERPA and Research with Human Subjects (Word document)
- Images and Recording in Human Subjects Research (Word document)
- Identifiable Datasets (Word document)
- Collection and Reporting of Demographic Information (Word document)
- Request for a "Not Human Subjects Determination" (opens in a new window)
    - Use this request form if humans are involved in your project but you are "unsure" whether you need IRB approval for your research involving AI.

Practical Considerations for Research with Human Subjects Using AI

- All AI tools, including existing tools with added AI capabilities, must undergo a security assurance assessment through the NC State University processes prior to being used, regardless of the data classification level. The appropriate data steward(s) must approve the use of the tool for the proposed use case. For some AI, this means going through the clickwrap process (opens in a new window) (e.g. freeware/shareware solutions that have an agreement to sign) and in others this means going through the IT purchase compliance process (ITPC) (opens in a new window).

- Considerations for all AI tool usage: Where is the data going? Where is the data stored? Where is the data processed? When the data gets to where it goes, what is the use agreement for that data?
    - The IRB is particularly concerned with cloud-based GAI because of how the data travels and is incorporated into the AI model, potentially increasing the exposure of that data to outside parties that may violate confidentiality expectations, participant privacy, or increase risks to participants from being in the study.

- For any research with human subjects where GAI or AI is used, the IRB performs a risk assessment for the study which includes a review of the "data classification levels" and the required data security measures for the use of a GAI or AI tool.
    - Though it is expected that all NC State researchers adhere to the "clickwrap" and "ITPC" processes referred to above and required by NC State, when using AI or GAI with "red" or "purple" data, the IRB will be more exacting in their review of protocols, ensuring the use of NC State licensed software for all non-exempt research or research that qualifies for an exemption with a limited review.
    - Any research that uses or generates "red" or "purple" data will require a data and access security plan (DASP) be submitted with their IRB protocol.

- **Participant Interactions**
    - **Recruitment and Informed Consent Processes**
        - Check to see that the AI models used for recruitment don't perpetuate or introduce biases (e.g., demographic, socio-economic, or psychological biases).

- Human intervention should occur to check that the "recruitment net" cast captures the intended participant groups.
- For AI models that are generally available, human intervention should occur to check AI for bias where possible.

- Be mindful of AI systems that could lead to participant manipulation (e.g., the creation of overly persuasive messages or targeting vulnerable populations). Either do not use them, intervene where possible, or adjust the study design.

- When consent is used, ensure the intent for current and possible future use is well described.
  - The NC State University IRB suggests adding broad consent (Word document) for future unspecified use of identifiable data in addition to the informed consent form and process.
  - In all consent materials and forms, clearly articulate "what we don't know" and the unknown about AI to ensure full transparency regarding the limits of current knowledge.

- When AI is used for translating informed consent, researchers must follow NC State University's IRB's verification of foreign translation process (Word document).
  - Confirmation of an AI translation must be verified via a human translator in alignment with the translation verifications guidance linked above.

- When GAI acts as an extension or representative of the investigator(s) by answering questions for potential, current, or past human research participants:
  - The AI tool should provide only factual and verifiable information.
  - The IRB does not permit AI tools to obtain automatic informed consent from participants; that is, a human investigator who is listed on the study and has completed all relevant training(s) must be present unless the study was approved with a waiver of consent or was determined to be exempt.
  - As with all consent processes, it is the responsibility of the principal investigator to ensure that the appropriate staff are in place and that participants adequately understand the

study to provide ongoing and meaningful informed consent.

- AI systems should not coerce or overly influence consent. Participants should have a clear, easy option to opt-out or discontinue their participation at any time with the knowledge of how to do so readily provided.

- **Data Collection**
  - Participants must be made aware of when, how, and where they will be interacting with GAI including what information the AI is accessing about them or generating with them.
    - If disclosure is not possible because knowledge of the AI involvement will influence responses/behavior and thus negatively influence the research data, please refer to the NC State IRB's guidance on research involving deception/incomplete disclosure.
  - Consider how GAI may influence the participant's responses (e.g., through tone, structure, or responses that might lead participants toward certain answers).
  - Ensure a clear mechanism is in place for participants to report concerns or request human intervention if necessary.
  - Consider if the GAI you are using or making is accessible to people with varying physical and intellectual abilities?
    - If no, be prepared to address inclusion/exclusion criteria with scientific justifications.
  - Do your participants have the intellectual capacity to understand the concept of GAI and its limitations, hallucinations, or that the GAI is not a human or other sentient being?
    - If the AI is interacting with participants, how will this need for assessment be addressed in study design?
    - If it is needed, how will capacity to understand be assessed or screened?
  - Are your participants in an age group that may need extra time or information to understand the needed information about the GAI?
    - How is this accounted for in the research design?

- **Participant Intervention Including Diagnosis, Treatment, or Other Behavioral Interventions**
  - When possible, participants and researchers should have full knowledge and control over any interventions (e.g., AI-driven recommendation systems or behavior nudges).

- Based on the IRB's risk assessment, investigators will be required to directly monitor any direct intervention between AI tools and participants. For example: Consider the psychological and physical risks of AI interventions (e.g., AI systems influencing mental health or behavior).

- For studies involving health related outcomes, a current best practice or standard of care must be "run at the same time" to corroborate the AI's recommendations/output.
  - In some cases, such as Food and Drug Administration (hereafter, FDA) regulated research with medical devices - this may be required
  - AI interventions must be ethically and legally permissible. If the AI involves a medical intervention (e.g., clinical decision support tools), it may require additional oversight and compliance with health regulations such as those from the FDA or HIPAA (Health Insurance Portability and Accountability Act).
  - AI interventions could inadvertently cause harm or unintended consequences (e.g., reinforcing stereotypes, creating dependence on technology).

- Datasets that are generated using AI can be biased because the data and algorithms used to create them can be biased. Researchers have greater control over this for homegrown/purpose-built AI than commercially available AI. Regardless of the AI model type, researchers must be mindful and account for:
  - Results in datasets that are not representative of the population of interest being studied and based on what the data are used for – this can cause direct harm to participants in the research.
  - Biases skewing the generalizability of research results, leading to invalid conclusions resulting in harm to historically marginalized and vulnerable populations by perpetuating negative stereotypes, discrimination, racism, classism, sexism, and other problematic assumptions.
  - Note: in cases where the study includes an interaction or intervention, risk assessments will determine the level of detail the IRB needs to know about the imputed data or how the AI was trained.

- **Manipulation of Participant's Environment**
  - Evaluate how AI-driven changes to a participant's environment may impact their behavior and results, and plan for contingencies. This will need to be communicated to the IRB.

- Minimize risks associated with environmental manipulation, including ensuring a safe and controlled environment. This will need to be communicated to the IRB.

- If AI is used to manipulate a participant's environment (e.g., through environmental triggers or virtual settings), this should be clearly explained to the participant, and their consent obtained unless deception/incomplete disclosure (Word document) is required. This may only occur in minimal risk research and IRB guidance should be adhered to.

- **Collection or Analysis of Directly Identifiable or Indirectly Identifiable Participant Data**
  - De-identification of human data is much more difficult with AI.
    - Anonymous, anonymized, or aggregated data sets may be more appropriate.
    - Data with identifiers are not appropriate for use with AI tools unless the participant has given explicit and prospective informed consent to do so, applicable laws and university regulations are accounted for, and the AI tool has undergone review in the ITPC process.

  - When providing personal information, NC State's Privacy and Security of Health Protected Information Regulation (opens in a new window) along with NC State University Data Management Regulation (opens in a new window) and IRB unit standards (opens in a new window) must be followed. There are restrictions around how data subject to FERPA, HIPAA (opens in a new window), the GDPR (PDF file), NIH's human genomic data (opens in a new window) can be used with AI, especially cloud based AI and GAI.

  - The National Institutes of Health (NIH) have noted that researchers who have been issued a Certificate of Confidentiality (CoC) for a study must consider the CoC protections when selecting third parties or vendors (including AI).
    - These third parties or entities (contractors, online platform vendors) must provide the same CoC protections including protecting CoC-covered information against compelled disclosure.

  - A GAI tool may not have access to the following data in identifiable/re-identifiable format, including the combination of indirect identifiers that could reasonably identify a participant.
    - Any data considered "red" or "purple" by NC State University
    - Biospecimens (urine, feces, saliva, blood, sweat, etc.)
    - Genomic data
    - Family Educational Rights and Privacy Act (FERPA) data
      - Some exceptions can be made for de-identified course artifacts and other data with direct permission from participants

        - HIPAA data (including limited datasets)
            - Some exceptions can be made for completely de-identified HIPAA data if HIPAA could consider the data anonymous and any relevant DUA's account for this use.
        - Data subject to the General Data Protection Regulation (GDPR) or California Consumer Privacy Act (CCPA) without direct permission from participants
        - Data that could reasonably put participants at risk such as audio/video recordings about sensitive or illegal topics.

    - Researchers using AI tools, including software and technologies owned by a third party, are responsible for understanding how the information provided to, or collected by these technologies, is protected and who has access to the data. Understanding the data flow is part of NC State University's IT purchase compliance process.

    - Refer to the NC State University IRB guidance on collecting and reporting demographics (Word document).

    - For cooperative research with partners beyond the NC State University community: Many academic medical centers have policies prohibiting entry of certain data types into public generative AI/LLM platforms including:
        - Protected Health Information (PHI) or individually identifiable information
        - Any portion of a medical record, including de-identified patient data
        - Proprietary business information, including research data
        - Other intellectual property owned by the institution or its employees

- **Data Generation and Sharing**
    - Researchers should retain control over the final interpretation of data rather than relying on the AI model to make conclusions about participants.

    - If you have permission from participants to put data into a GAI via an informed consent form reviewed and approved by the IRB and associated with an IRB approved protocol, you may do so.
        - Note: some software already integrates aspects of AI. If the software is licensed and vetted by NC State and the OIT data sensitivity framework requirements are followed, this use of AI is allowed if it isn't generating new information about participants nor is it being shared widely.

    - All federal and state laws apply to secondary data use, so be sure to know if the data you are going to use with AI is subject to HIPAA, FERPA, the GDPR (opens in a new window) and other privacy and security laws.

- The IRB recommends that investigators limit an AI tool's access to demographic information and other data points that could re-identify an individual (e.g., combinations of data points that could identify an individual).

- The IRB recommends and, in certain situations based on a risk assessment, will require that investigators place limitations and parameters on the AI or Generative AI tool's data use.

- When using data from a GAI that you did not input, remember that that data is likely unreliable. Data integrity is always the key factor when choosing to use a GAI tool to assist with a research project. During the planning phase, if this cannot be maintained and documented then it is not advisable to use.

- **Transcription**
  - The IRB considers AI tools providing transcription services to be integral to human research.

  - NC State's licensed Zoom's transcription service has been reviewed for security assurance and is approved by NC State for human subjects' data with the sensitivity levels of up to "red" - highly sensitive. Users of the NC State's Zoom environment must also be compliant with Rule 08.00.18 – Endpoint Protection Standard (opens in a new window). Researchers can set Zoom meeting security controls to meet HIPAA by using the Counselor Education template available in "Meeting Templates."
    - Using Zoom's AI Companion Tools at NC State (opens in a new window)
    - Overview of Zoom meeting templates (opens in a new window)

  - Fireflies AI (opens in a new window) has been reviewed for security assurance and is approved for privacy, confidentiality, and security by NC State for human subjects' data with the sensitivity levels of up to highly sensitive. Users must also be compliant with Rule 08.00.18 – Endpoint Protection Standard (opens in a new window). Note: NC State strongly recommends the use of enterprise Zoom rather than paying for tools that provide comparable features.

What to Include in the IRB Application

- A clear statement as to if the GAI or AI is commercially available/off the shelf or if the AI is made or modified at NC State. If downloaded and trained, additional statements to describe that will be needed.

- A clear statement of whether the GAI or AI will analyze and interact with data

locally (internal to the NC State network) or not.

- If possible, a clear statement as to how the GAI or AI is trained and what data it will be trained from (including what the human intervention in its training includes).
    - For GAI or AI where the NC State researchers are completing the AI's training and inputting the data, this must be detailed in the IRB application.

- A full description of the planned interaction, intervention, manipulation of environment between the participant and the GAI or AI tool (scripts and screenshots are encouraged).

- A description of the data that the GAI or AI tool will be designed to collect and an assessment of the data's identifiability (Word document) and the classification of the data (Google Sheet).

- Documentation of the parameters or limits placed on the GAI or AI tool used for the intervention, data collection, and (if applicable) data analysis. These can include cut off parameters, descriptions of what will not occur, boundaries set. The details should reflect the risk assessment.

- If possible, scripts or texts of instructions that will be read or provided to participants as part of the interaction with the GAI or AI tool.

- Disclosure if AI or GAI tools are being used to facilitate recruitment (e.g., automated email communications, chatbots for preliminary screening). Participants should know how AI tools impact their recruitment process.

- If appropriate to the risk assessment, a plan to monitor the safety of participants and their data during and after an intervention.

- An information sheet that the participant can refer to about accessing, installing, using, settings, and removal of the GAI or AI (as appropriate to the study).
    - Refer to the IRB's applications and software in human subjects research guidance (Word document) for more information.

- Information regarding if individual results/findings will be shared/returned to the participants, especially as related to health-related interventions or assessment.

- If appropriate to the risk assessment, an already existing reliable intervention should be run concurrently to ensure that the GAI or AI output/recommendations are in line with best practices.
    - This will likely be required for studies undergoing convened full board review.
    - FDA regulated research requires that a standard of practice process/device/treatment be concurrently run (FDA IDE Regulations –

opens in a new window).

- If the GAI or AI is being used to diagnose, treat, cure, or prevent disease or other conditions, please refer to the medical devices unit standard (Word document) as clinical algorithms are regulated as medical devices.

- Description of the current phase of research (Word document) in the specific protocol/IRB application and how the AI model is functioning within that phase:
    - Proof-of-concept, developing, training, validating, or testing a model - not impacting care
    - Real-world clinical trial, validating or testing a model with an impact on care
    - Algorithms
        - Does the algorithm inform or drive clinical decision-making?
            - What does the clinical back-up and decision-making look like?
        - Does the algorithm diagnose or treat?
            - Are decisions autonomously acted on?
        - Does the algorithm change over time?
        - How is the algorithm expected to work and how does it make its decisions?
        - What are the anticipated ways in which the algorithm may make mistakes?

- A statement that training datasets are adequately representative and if not, why. If this is not possible due to the AI tool selected for use, this must be clearly articulated.

- A statement about how continuous monitoring will occur for unanticipated problems and bias.

- A statement regarding the inclusion of community/stakeholder input, including practitioners or representatives of the target population, etc.

- When applicable to what is being studied, how will the research design determine the participant's ability to tell the difference between their interactions with humans and their interactions with GAI or AI - including what is "real," what is a "hallucination," or "poor training."
    - This can be done through assessing participants' capacity for participation.
    - This can be mitigated through participant screening and selection, including prospective disclosure and throughout the intervention/interaction, or as a debriefing.

What to Disclose in the Informed Consent Form

- Use the NC State University IRB informed consent templates (opens in a new

window) to begin and adapt it for your research activities.

- Disclose how GAI or AI is used, e.g., generating content, making recommendations, making decisions, and how it works in lay terms.

- Disclose if the GAI or AI is off the shelf, commercially available, downloadable through an app store, made somewhere else, or homegrown/purpose-built. Though there are many gray areas between these statuses, this disclosure is important for transparency purposes and participant autonomy around risk assessment. For example, NC State University has more control over purpose-built, “homegrown” AI than others.

- Disclose if the use of GAI or AI is being used in an experimental way to address health related outcomes.

- Disclose how the GAI or AI was trained, e.g., representative datasets and if this is not possible, state that it is not known.

- Disclose anticipated risks specific to the GAI or AI.

- Acknowledge unanticipated risks with GAI or AI (e.g., emergent behavior) using the appropriate NC State University IRB informed consent template.

- Describe how data will be used – how data from the present study will be used by the GAI or AI in this study and in future uses.

- Describe any limitations of personal privacy and data confidentiality, e.g., related to third-party software use.

- Describe potential risks related to re-identification, including in future uses (e.g., commercialization or open-source use) specific to the use of the GAI or AI tool.

- Information regarding if individual results, recommendations, or findings from the use of the GAI or AI tool will be shared/returned to the participants.

- Describe how the participants data is safeguarded in transit and rest and where it will be stored (cloud/on premise solution, NC State hosted, vendor, or vendor’s third-party provider.

- Describe whether the participants data will be collected by a third-party vendor for training GAI or AI modules and what data will be collected.

- Link the security and privacy statements for the GAI or AI tool.

Guidance from Federal Agencies

- SACHRP Committee (2022, October 19). *IRB Considerations on the Use of Artificial Intelligence in Human Subjects Research*. U.S. Department of Health and Human Services. https://www.hhs.gov/ohrp/sachrp-committee/recommendations/irb-considerations-use-artificial-intelligence-human-subjects-research/index.html (opens in a new window)

- National Science Foundation. (2023, October 4). *Artificial intelligence*. National Science Foundation. https://www.nsf.gov/focus-areas/artificial-intelligence (opens in a new window)

    - National Science Foundation. (2023, August 1). *National Artificial Intelligence Research Institutes*. National Science Foundation. https://www.nsf.gov/funding/opportunities/national-artificial-intelligence-research-institutes (opens in a new window)

- National Institutes of Health, Office of Science Policy. (n.d.). *Artificial intelligence*. U.S. Department of Health and Human Services. https://osp.od.nih.gov/policies/artificial-intelligence/ (opens in a new window)

    - National Institutes of Health, Office of Data Science Strategy. (2025, March 11). *Artificial intelligence at NIH*. U.S. Department of Health and Human Services. https://datascience.nih.gov/artificial-intelligence (opens in a new window)

    - National Institutes of Health, Office of the Director. (2025, March 28). *Protecting human genomic data when developing generative artificial intelligence tools and applications (Notice No. NOT-OD-25-081)*. U.S. Department of Health and Human Services. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-25-081.html (opens in a new window)

- U.S. Department of Defense. (2020, February 24). *DOD adopts ethical principles for artificial intelligence*. https://www.defense.gov/News/Releases/release/article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/ (opens in a new window)

    - U.S. Department of Defense. (2021, May 26). *Implementing responsible artificial intelligence in the Department of Defense.* https://media.defense.gov/2021/May/27/2002730593/-1/-1/0/IMPLEMENTING-RESPONSIBLE-ARTIFICIAL-INTELLIGENCE-IN-THE-DEPARTMENT-OF-DEFENSE.PDF (PDF file)

- U.S. Food and Drug Administration. (n.d.). *Artificial intelligence and machine learning in software as a medical device.* U.S. Department of Health and Human

Services. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device (opens in a new window)

- U.S. Food and Drug Administration. (n.d.). *Artificial intelligence and machine learning (AI/ML)-enabled medical devices*. U.S. Department of Health and Human Services. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices (opens in a new window)

- U.S. Food and Drug Administration. (2024, June 13). *Transparency for machine learning-enabled medical devices: Guiding principle*s. U.S. Department of Health and Human Services. https://www.fda.gov/medical-devices/software-medical-device-samd/transparency-machine-learning-enabled-medical-devices-guiding-principles (opens in a new window)

Examples of AI Research Requiring IRB Approval

A researcher wants to create a GAI that can teach people how to properly lift weight. The researcher is at the beginning stages and needs video recordings of humans lifting items so that the AI has data to pull from to start training.

- The videos are from publicly available sources and there are no restrictions on use. This DOES NOT need IRB approval because the data are publicly available and you are not interacting, intervening, or manipulating a participant's environment.
- The researcher requests that participants come to a place and lift items while being video recorded. This DOES need IRB approval because you are interacting/intervening with a human for research purposes and the data you are generating to use is about their body/skills.

A researcher wants to train an already existing and publicly available AI to effectively assist teachers in grading student papers.

- The researcher wants to use their own students' coursework submitted as a part of their classes and put that coursework into the AI to compare how the AI graded the work versus how they themselves graded the work. This DOES need IRB approval because the researcher is using private FERPA records for research purposes.
- The researcher finds example assignments that are publicly available to put into the AI to compare how the AI graded the work versus their own analysis of what the example assignment would warrant. This DOES NOT need IRB approval because the researcher is using publicly available assignments that are either fake, de-identified, or identifiable but public.

A researcher wants to use AI to create a “case overview” of individual clients/students/patients experiences to share with colleagues for use in providing overviews of, or services/individualized curriculum/individualized treatment, for people they work with. Refer to below examples for varying contextual information.

- The researcher is a licensed private therapist with their own private business, and they are also a faculty member at NC State. They want to access and use individual patient HIPAA covered information from their private practice for research purposes in their role as an NC State faculty member. They want to put the patient’s information into a Generative AI to have the AI assess the information and provide insight into how to move forward with treatment ideas. The researcher does not want to implement the treatment, but they do want to analyze if the AI generated “standard” and “good” treatment recommendations for the individuals. This DOES need IRB approval because the data used is about the individual.
    - This likely needs additional agreements related to the researcher’s role (such as an individual investigator agreement).
    - This will need additional authorization from individual participants to address release of HIPAA records used in research with AI. The IRB will need to consider the records imputed and the AI’s ability to handle the regulatory requirement for HIPAA data.

- The researcher is a faculty member at NC State and they work at the vet school. They want to use individual information from the vet’s school’s clients (animals) cases that includes information about the decision making from the humans who own the animals. The faculty member within the scope of their role as an NC State faculty/vet staff, wants to put all information from the client’s case into an AI which will then summarize the issues succinctly to assist other veterinary professionals in education, training, or practice.
    - This does need IRB approval if the information imputed into the AI has direct or indirect identifiers and has information about the human’s decision-making rationale or considerations
    - This DOES NOT need IRB approval if there are no direct or indirect identifiers that could lead to re-identification of an individual human.

- The researcher is a student at NC State completing their dissertation. They are also a public-school teacher at a local elementary school in NC. As a part of their work in school they are using Generative AI to assist with addressing students’ behavioral issues - both attempting to predict situations that may trigger individual students' negative behaviors and recommend interventions to put in place before the triggering event occurs. They think that this has some really great potential and want to research it for their dissertation. For research purposes, the NC State student would like to access students’ grades, behavioral reports, and other documentation about the student and enter the information into a generative AI to predict when they could "expect" to see behavioral issues manifest and to provide appropriate interventions that may intervene and prevent those behavioral issues from occurring.

    - This DOES need IRB approval if any of the data accessed is directly or indirectly identifiable due to triangulation of data points or sources or researcher access and expertise.
    - This DOES NOT need IRB approval if the data are completely anonymized by someone not on the research team and the researchers and AI have no way to re-identify an individual from the data or its analysis. Note, this is not possible for a researcher to do using their own records or records they have access to due to their role. Note, other laws or regulations outside of the IRB regulations may apply (such as FERPA or other university regulations)

- A researcher wants to create and train an AI to diagnose and provide recommendations for treatment for a medical issue related to how a person physically moves.
    - This DOES need IRB approval as it is considered testing the safety and efficacy of a medical device and is subject to the IRB regulations and the FDA.
    - Note, these types of studies have MANY stages/phases. There is a chance that in some of the initial stages/phases of the study, when working on the AI itself without interacting, intervening, using private identifiable data, or manipulating the participants' environment, that that aspect does not need IRB approval. Please reach out to the NC State IRB for questions or confirmation regarding specific non-hypothetical cases.
    - Note, this could also apply to medical issues around animals – but the content of the data will influence whether or not it needs IRB approval.

- A researcher wants to use AI to assist in data analysis of quantitative and qualitative data for research purposes.
    - This DOES need IRB approval if the data are identifiable or re-identifiable to the researcher or the AI. This includes direct and indirect IDs including unique stories or experiences.
    - This MAY need IRB approval based on the type of AI used (generative or not), how the AI handles the data, and if the AI or the researcher can possibly re-identify an individual from the imputed data or its analysis.
    - This DOES NOT need IRB approval if there are no direct or indirect identifiers on the data imputed into the AI and research team and the AI cannot re-identify an individual from the imputed data (through triangulation of other data points or sources, or researcher/AI expertise, access, or role). Though other laws may apply that would prohibit use of the data in this way, the IRB does not have jurisdiction over the use of completely anonymous human subjects' data.

- A researcher wants to use AI to generate an outline of study design with suggestions for recruitment, consent language, and modes of data collection to address the posed research questions and to create an instrument that will be used to collect data from humans

- This DOES need IRB approval if the study design and material are implemented with participants for research purposes. Should this occur, minimal information about the AI use would be needed in the IRB protocol as it is not interacting, intervening, accessing private identifiable data about, or manipulating the environment of participants.
- This activity DOES NOT need IRB review and approval because of how the AI is used only to create and design the study.

Appropriate Data/Datasets for Use with Cloud-Based AI

This section is intended to provide guidance on how researchers can clean and prepare human subjects' research datasets for analysis using AI or GAI. No data considered "red" or "purple" can be imputed into a cloud-based AI.

Note: the information in this section may change before the guidance is updated. The rate of change is far faster than the guidance updates every few months.

- **Datasets Subject to FERPA**
  - If applicable, review the data use agreement (DUA) associated with the data intended for analysis with cloud-based AI or GAI and ensure that the data are allowed to be shared in any way.
    - If it cannot be shared as stated in the DUA, it cannot be imputed into a cloud-based AI.
  - For research with human subjects, review the IRB protocol associated with the data intended for analysis with cloud-based AI or GAI and ensure that the data are allowed to be shared in any way.
    - If it cannot be shared as stated in the IRB application AND informed consent form AND FERPA records release or permission, it cannot be imputed into a cloud-based AI.
  - If the data are released to the researcher as a dataset from an organization such as Registration and Records at NC State:
    - If there is a university policy or regulation in place that indicates the FERPA data cannot be imputed into cloud-based AI, then that must be followed.
  - If the data are accessed through your own (or a colleagues) access to course records, then:
    - If there is a university policy or regulation in place that indicates the FERPA data cannot be imputed into cloud-based AI, then that must be followed.
    - All direct or indirect identifiers MUST be removed before any data are imputed into the cloud-based AI.
  - In general - the following data categories related to FERPA data CANNOT be imputed into cloud-based AI:

        - Any direct or indirect participant identifier
            - Direct Identifiers: Direct identifiers include information that relates specifically to an individual such as the individual's residence, including for example, name, address, Social Security number (SSN) or other identifying number or code, telephone number, e-mail address, or biometric record.
            - Indirect Identifiers: information that can be combined with other information to identify specific individuals, including, for example, a combination of gender, birth date, geographic indicator and other descriptors. Other examples of indirect identifiers include place of birth, race, religion, weight, activities, employment information, medical information, education information, and financial information.
        - Any FERPA data categories listed as "red" or "purple" in the NC State Data Classification Table (Google Sheet)
        - Any directly or indirectly identifiable data: Refer to NC State University IRB's identifiable/re-identifiable datasets guidance (Word document)

- **Datasets Subject to HIPAA**
    - The NC State IRB HIPAA unit standard (Word document) must be adhered to.

    - If applicable, review the DUA associated with the data intended for analysis with AI or Generative AI and ensure that the data are allowed to be shared in any way.
        - If it cannot be shared as stated in the DUA, it cannot be imputed into a cloud-based AI.

    - If applicable, review the IRB protocol associated with the data intended for analysis with AI or Generative AI and ensure that the data are allowed to be shared in any way.
        - If it cannot be shared as stated in the IRB application AND informed consent form AND the HIPAA authorization, it cannot be imputed into a cloud-based AI.

    - If the data are released to the researcher as a dataset from an organization such as Student Health at NC State, the data must be cleaned in its entirety with no direct or indirect identifiers, and it should be able to be considered "green" data before being imputed into a cloud-based AI.

    - If the data are accessed through your own access to NC State medical records, you must adhere to the NC State IRB unit standard for HIPAA data and the data must be cleaned in its entirety with no direct or indirect

identifiers and it should be able to be considered “green” data before being imputed into a cloud-based AI.

- In general - the following data categories related to HIPAA data CANNOT be imputed into cloud-based AI:
  - Any directly or indirectly identifiable data: Refer to the identifiable/re-identifiable data guidance (Word document) from the IRB
  - Any HIPAA data categories listed as “red” or “purple” in the NC State Data Classification Table (Google Sheet)
  - Name
  - Address (all geographic subdivisions smaller than state, including street address, city county, and zip code)
  - All elements (except years) of dates related to an individual (including birthdate, admission date, discharge date, date of death, and exact age if over 89)
  - Telephone numbers
  - Fax number
  - Email address
  - Social Security number
  - Medical record number
  - Health plan beneficiary number
  - Account number
  - Certificate or license number
  - Vehicle identifiers and serial numbers, including license plate numbers
  - Device identifiers and serial numbers
  - Web URL
  - Internet Protocol (IP) Address
  - Finger or voice print
  - Photographic image - Photographic images are not limited to images of the face.
  - Any other characteristic that could uniquely identify the individual

- **Datasets Subject to the GDPR**
  - The IRB recommends that NC State researchers seek broad consent (opens in a new window) for the future unspecified research use of identifiable data.
    - Broad consent is sought in addition to informed consent in human subjects research when the researcher implements primary data collection.

  - If applicable, review the data use agreement (DUA associated with the data intended for analysis with any cloud-based AI or GAI and ensure that the data are allowed to be shared in any way.

        - If it cannot be shared as stated in the DUA, it cannot be imputed into a cloud-based AI.

    - If applicable, review the IRB protocol associated with the data intended for analysis with cloud-based AI or GAI and ensure that the data are allowed to be shared in any way.
        - If it cannot be shared as stated in the IRB application AND informed consent form and, if applicable, the broad consent form, it cannot be imputed into a cloud-based AI.

    - In general - the following data categories related to the GDPR CANNOT be imputed into cloud-based AI:
        - Ethnic or racial origin
        - Political opinions
        - Cultural or social identity
        - Philosophical beliefs
        - Religious beliefs
        - Trade union memberships
        - Genetic data
        - Biometric data (that uniquely identifies someone)
        - Name and surname
        - Email address
        - Phone number
        - Home address
        - Date of birth
        - Race
        - Gender
        - Political opinions
        - Credit card numbers
        - Data held by a hospital or doctor
        - Photograph where an individual is identifiable
        - Identification card number
        - A cookie ID
        - Internet Protocol (IP) address
        - Location data (for example, the location data from a mobile phone)
        - The advertising identifier of a phone (a unique string of characters assigned to a device for advertising purposes)

- **General Datasets**
    - If applicable, review the data use agreement (DUA) associated with the data intended for analysis with cloud-based AI or Generative AI and ensure that the data are allowed to be shared in any way.
        - If it cannot be shared as stated in the DUA, it cannot be imputed into a cloud-based AI.

- Review all IRB protocols associated with the data intended for analysis with cloud-based AI or Generative AI and ensure that the data are allowed to be shared in any way.
    - If it cannot be shared as stated in the IRB application AND informed consent form AND broad consent form, it cannot be imputed into a cloud-based AI.

- In general - the following data categories related to human subjects' research data CANNOT be imputed into cloud-based AI:
    - Any data categories listed as "red" or "purple" in the NC State Data Classification Table (Google Sheet)
    - Any directly or indirectly identifiable data: Refer to identifiable/re-identifiable data guidance (Word document) from the NC State University IRB

Appendix A
Roles and Responsibilities

Principal Investigator (PI)
The PI is responsible for understanding the AI tools they use in their research as well as complying with all university policies and applicable state, federal, and international regulations, including those dealing with copyright and other intellectual property.

Research Team Members
Research team members who generate, acquire, and work with research data are responsible for understanding the AI tools they use in their research as well as complying with all university policies and applicable state, federal, and international regulations, including those dealing with copyright and other intellectual property.

Office of Information Technology (OIT)
OIT is tasked with staying abreast of advances, novel threats, and emerging new security-related tools related to AI and providing security training, security tools, consultation, and guidance to researchers and department IT personnel in the changing landscape.

University Departments
Departments are responsible for regularly analyzing risks related to their technology assets, including the integration of AI tools.

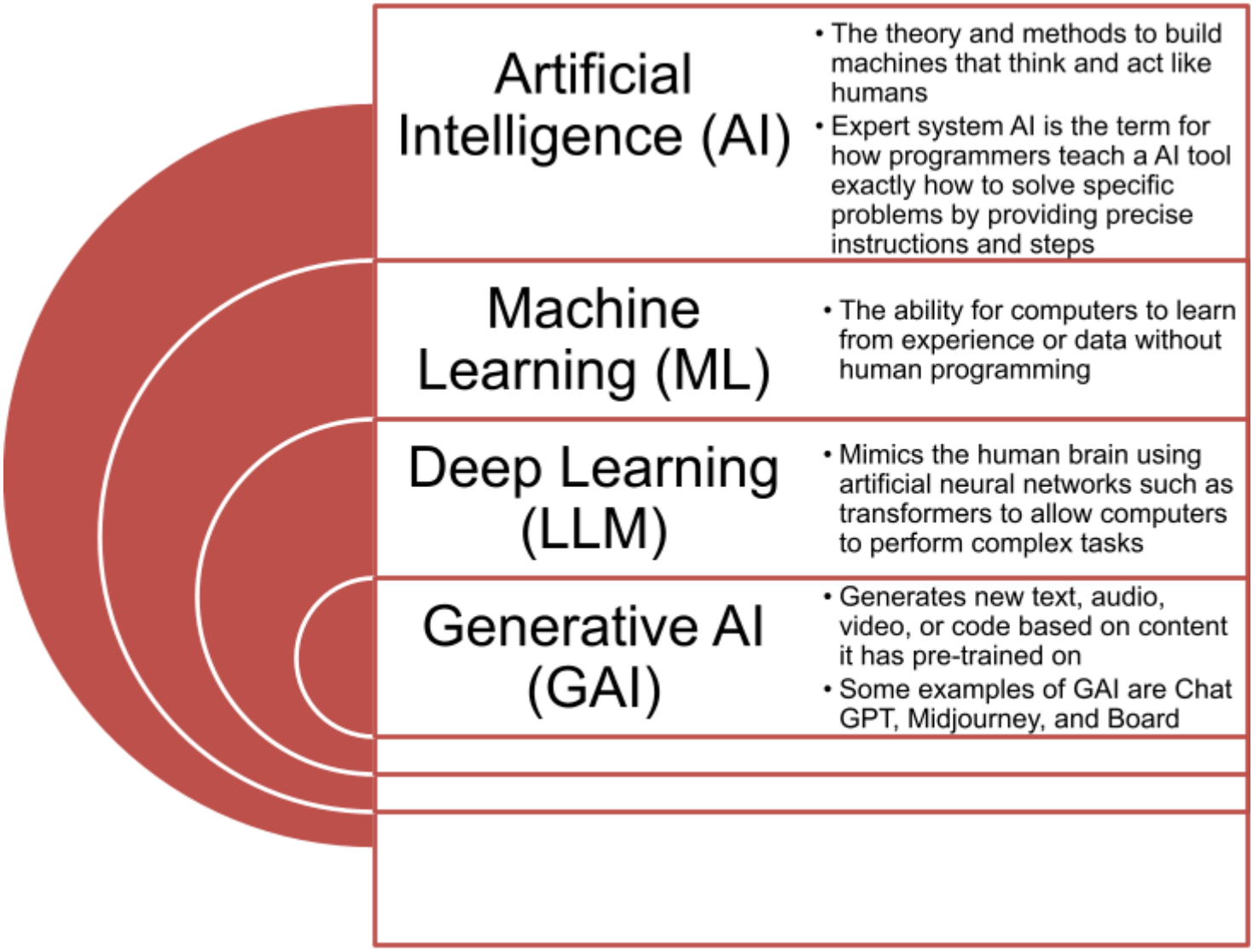
Research Compliance
Research compliance is responsible for providing guidance to researchers for implementing the appropriate confidentiality, privacy, and security protections when using AI tools to collect and process research data.

IRB Office
The IRB office is responsible for supporting researchers who incorporate AI tools in their research with meeting their ethical and regulatory responsibilities to human research participants.

Appendix B
Resources

Ampersand (2023, August 1). *Artificial Intelligence (AI) Is Moving Fast – What Are the Ethical Implications for Scientific Research?* Public Responsibility in Medicine and Research (PRIM&R). https://blog.primr.org/artificial-intelligence-ai-is-moving-fast-what-are-the-ethical-implications-for-scientific-research/ (opens in a new window)

Fjeld, J., Achten, N., Hilligoss, H., Nagy, A. & Srikumar, M. (2020). *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI.* Berkman Klein Center for Internet and Society at Harvard University. https://dash.harvard.edu/bitstream/handle/1/42160420/HLS%20White%20Paper%20Final_v3.pdf?sequence=1&isAllowed=y - PDF (PDF file)



Above image is adapted from AI for Education. (2024, November 29). *Generative AI explainer*. AI for Education. https://www.aiforeducation.io/ai-resources/generative-ai-explainer (opens in a new window)