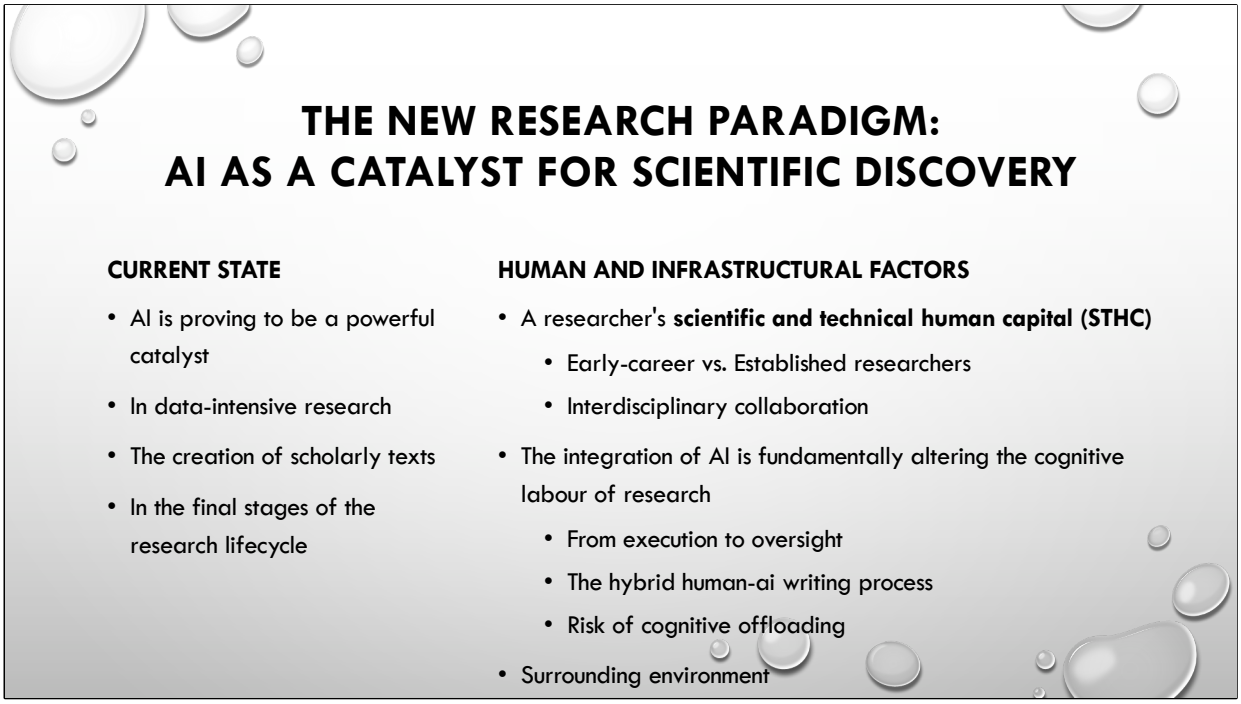
NAVIGATING THE INTEGRATION OF ARTIFICIAL INTELLIGENCE IN SCIENTIFIC RESEARCH

MÁRTA TURCSÁNYI-SZABÓ,

habil. associate professor
Dean's commissionaire for innovation in education



The rapid proliferation of Generative Artificial Intelligence (AI) represents a fundamental and enduring shift in the scientific process, not a fleeting trend. This paradigm shift presents both unprecedented opportunities for innovation and significant challenges that demand strategic navigation from the entire research enterprise. Since 2015, scientific research has entered a new renaissance, powered by a confluence of advances in computational power, an abundance of data, and significant algorithmic innovation. Data reveals a sharp, cross-disciplinary increase in the integration of AI methods into scholarly work, moving well beyond computer science to impact several fields

Artificial intelligence is being deployed not as a niche technology for specific problems but as a multifaceted assistant across the entire spectrum of research activities. From the initial spark of an idea to the final dissemination of findings, AI tools are streamlining workflows, augmenting cognitive tasks, and creating new efficiencies. An analysis of its current applications reveals a pervasive integration at every stage.

AI is proving to be a powerful catalyst in the earliest, most creative phases of research. It is used to brainstorm novel ideas, generate testable hypotheses, and

rapidly scan vast bodies of literature to identify critical research gaps. Furthermore, researchers are leveraging AI to draft initial designs for complex experimental studies, construct robust survey questionnaires, and even generate preliminary methodological frameworks for review, significantly accelerating the foundational work of a project.

In data-intensive research, AI is streamlining tasks that were once major bottlenecks. It is increasingly used for automating the cleaning of messy data, harmonizing heterogeneous datasets from disparate sources, and performing sophisticated exploratory analyses to uncover hidden patterns and correlations that might elude human observation. By applying advanced statistical tests, AI is unlocking deeper insights from existing data.

The creation of scholarly texts has been profoundly impacted by AI. Its functions range from drafting initial sections of a manuscript and summarizing complex articles to improving the clarity, coherence, and style of a researcher's writing. This capability has been particularly crucial in providing language support, helping non-native English speakers overcome linguistic barriers and participate more equitably in the global academic discourse.

In the final stages of the research lifecycle, AI is emerging as a valuable tool for navigating the complexities of academic publishing. It can help researchers identify suitable journals for their manuscripts, assist in preparing structured and well-reasoned rebuttals to peer reviewer comments, and adapt dense, jargon-filled academic work into accessible summaries for broader public engagement and science communication.

The successful and equitable integration of AI into research is shaped less by the technology itself and more by a critical combination of human, social, and infrastructural variables. Evidence suggests that adoption patterns are primarily driven by the capabilities of researchers and the culture of their institutions, rather than sheer technological availability. Understanding these factors is essential for fostering an environment where AI can be leveraged effectively and responsibly.

A researcher's **Scientific and Technical Human Capital (STHC)**—defined by their pre-existing skills, collaborative networks, and intrinsic "taste for exploration"—is the primary driver of AI adoption. This human element proves to be a stronger predictor than institutional resources or demographic characteristics.

- **Early-Career vs. Established Researchers:** Adoption patterns differ across career stages. Early-career researchers ("newbies"), who often possess more current AI skills, are frequently instrumental in bringing AI techniques into research teams. Conversely, established scientists with a high citation stock may be more hesitant to adopt new methods, potentially perceiving a "risky gamble" in deviating from proven

approaches.

Interdisciplinary Collaboration: Critical AI skill gaps are often bridged through collaboration. The data shows a growing reliance on interdisciplinary teams that pair domain experts with computer scientists, compensating for a lack of specialized AI knowledge within a given field.

The integration of AI is fundamentally altering the cognitive labor of research, shifting the primary effort from direct task execution to oversight and management.

From Execution to Oversight: Researchers are spending less time on tasks like writing first drafts or running basic analyses and more time on prompt engineering, critical verification of AI outputs, and sophisticated editing. This creates new cognitive demands centered on evaluation and refinement.

The Hybrid Human-AI Writing Process: Effective use of AI in writing has been observed as a dynamic and continuous "negotiation of meaning." In this model, the researcher and AI act as collaborative partners, a process that can enhance the researcher's agency, confidence, and cognitive engagement.

Risk of Cognitive Offloading: A documented risk is the tendency toward cognitive offloading or over-reliance on AI. If left unmanaged, this can lead to the atrophy of independent problem-solving skills and a reduction in critical reflection, turning a powerful tool into a cognitive crutch.

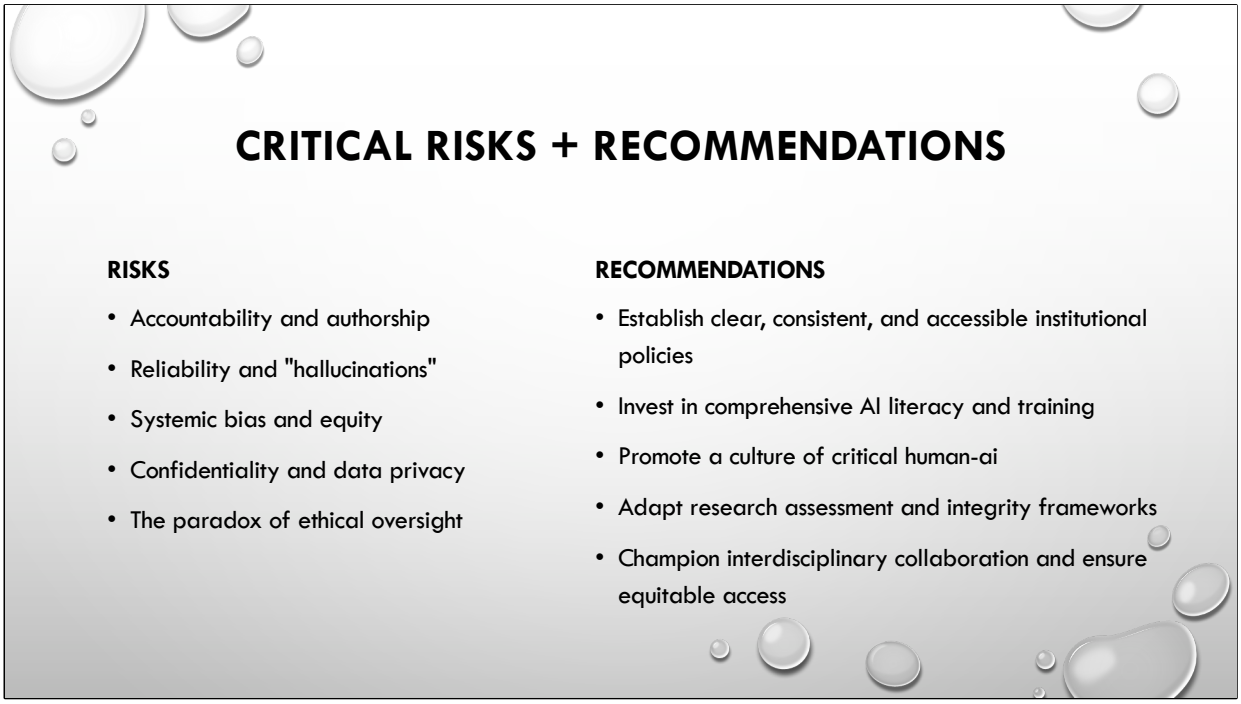
While technology is not the sole driver, the surrounding environment plays a crucial role in either enabling or hindering AI adoption.

The "Compute Divide": Access to infrastructure like High-Performance Computing (HPC) is a critical factor for AI adoption, but only in a few specific, data-intensive fields such as genomics or particle physics. For most other disciplines, this "compute divide" is not a primary barrier.

Institutional Environment: The local research culture has a profound influence. An institution's specialization in AI and the visible presence of peer adopters can significantly increase an individual researcher's likelihood of integrating AI into their own work.

Ambiguous Institutional Policies: A significant hurdle is the current state of institutional governance. Policies regarding AI use are often inconsistent, ambiguous, or difficult for researchers to locate. This creates uncertainty and hesitation rather than providing the clear, supportive guidance needed for responsible innovation.

Ultimately, successful AI integration depends less on providing raw access to technology and more on a concerted strategy to cultivate human capital. Without clear, supportive institutional frameworks, the cognitive risks of over-reliance are unmanaged, skill gaps are bridged only by ad-hoc collaborations



Navigating Critical Risks and Ethical Imperatives

The transformative potential of AI is matched by profound ethical and practical challenges that research institutions and policymakers must urgently address. Failure to proactively manage these risks threatens to undermine scholarly integrity, erode public trust in science, and amplify existing societal inequities.

Accountability and Authorship: There is a clear and established consensus among international funding bodies, such as the European Commission, and major publishers: AI systems cannot be designated as authors. Authorship fundamentally implies responsibility and accountability for the work, attributes that must remain exclusively with human researchers.

Reliability and "Hallucinations": AI models are prone to "hallucination," a phenomenon where the system generates confident-sounding but entirely fabricated information. This can manifest as fictitious citations, misrepresented data, or invented facts. The inherent unreliability of AI-generated content for factual consistency makes rigorous, meticulous human verification a non-negotiable imperative.

Systemic Bias and Equity: AI models are trained on vast datasets that reflect existing societal and historical biases. This creates a significant risk that these systems will not only perpetuate but actively amplify these biases. Evidence

already suggests that the benefits of AI in research are not being distributed equitably, with research showing that women and underrepresented minority (URM) scientists appear to profit less from engaging with AI compared to their male and non-URM counterparts.

Confidentiality and Data Privacy: Inputting sensitive, confidential, or unpublished data—such as manuscripts under peer review or human subject data—into public AI platforms constitutes a major security risk. Publishers and funding agencies have issued explicit warnings that this practice can lead to serious breaches of confidentiality, as the data may be stored or used for future model training.

The Paradox of Ethical Oversight: The responsible use of AI in research introduces a critical paradox. To ensure ethical compliance and scholarly rigor, researchers must conduct thorough human oversight and verification of all AI-generated outputs. This necessity often requires them to perform the very tasks they intended to outsource (e.g., reading original sources to verify an AI-generated summary). This paradox challenges, and in some cases completely negates, the promised efficiency gains that motivate AI adoption in the first place.

Proactive and decisive governance is essential to mitigate these risks before they erode the foundational principles of the scientific enterprise.

5.0 Recommendations for Fostering a Responsible AI-Enhanced Research Ecosystem

To harness AI's transformative power while upholding the core values of science, academic administrators and research policymakers must adopt a strategic and principled framework. The following recommendations are designed to guide the responsible, ethical, and productive integration of AI into the research ecosystem.

Establish Clear, Consistent, and Accessible Institutional Policies Institutions must develop and prominently disseminate clear, university-wide guidelines for the use of generative AI in research. These policies must, at a minimum, mandate **transparency** through the explicit disclosure of AI use and affirm **accountability** by placing absolute responsibility for all research outputs on human authors. These standards should align with emerging best practices from international publishers and funding bodies.

Invest in Comprehensive AI Literacy and Training Robust, ongoing training programs are essential for students, faculty, and research staff. This education must move beyond basic tool usage to build advanced skills in prompt engineering, the critical evaluation of AI-generated content, and a deep, nuanced understanding of AI's limitations, including its potential for bias, error, and hallucination.

Promote a Culture of Critical Human-AI Partnership Foster a research and pedagogical culture that frames AI as a collaborative partner designed to

augment, not replace, human intellect. Encourage practices that involve a "negotiation of meaning" with AI systems, promoting active cognitive engagement and preventing the development of academic indolence or uncritical over-reliance.

Adapt Research Assessment and Integrity Frameworks Formally re-evaluate and update institutional standards for academic integrity, research assessment, and peer review to explicitly account for AI's role. This includes updating definitions of plagiarism to address "AIgiarism" (the unethical use of AI to generate work presented as one's own) and providing clear guidance to reviewers and evaluation committees on how to assess research created with AI assistance.

Champion Interdisciplinary Collaboration and Ensure Equitable Access Allocate funding and create structural support for interdisciplinary teams that unite domain experts with AI specialists. Furthermore, institutions must actively address the documented equity gaps in AI adoption by creating targeted programs, mentorship opportunities, and resource access to ensure researchers from underrepresented groups can fully and equitably benefit from AI's potential.

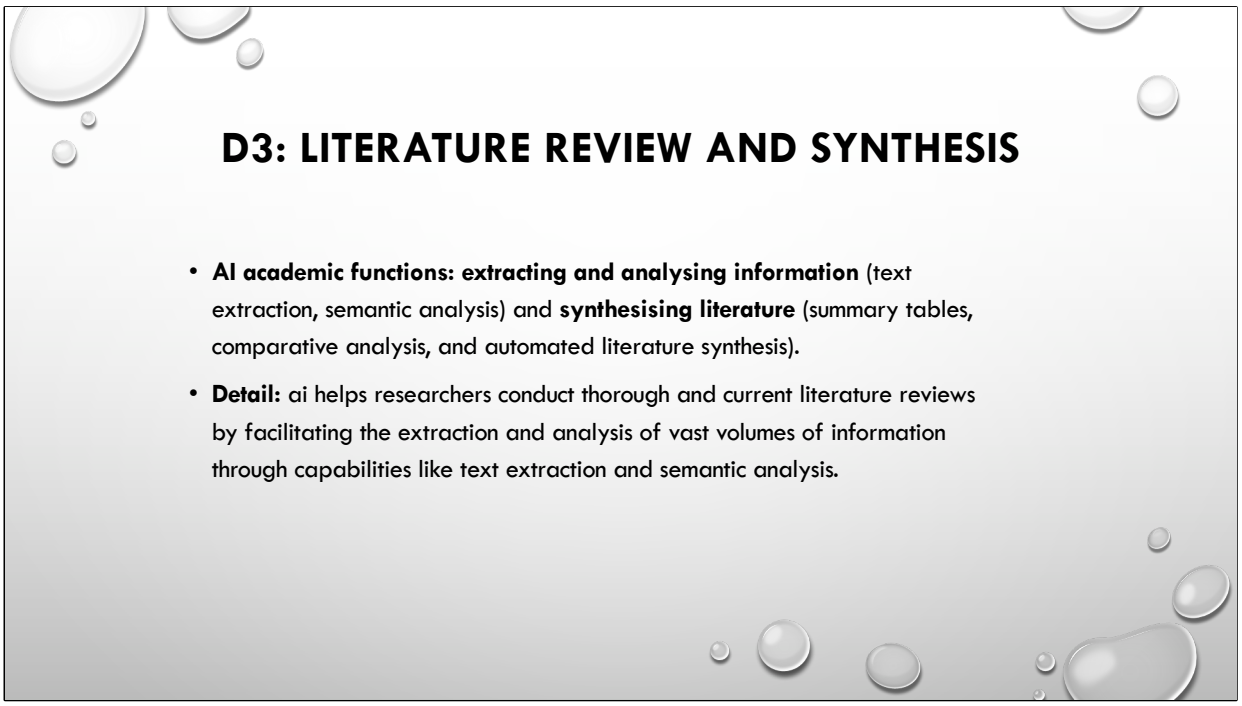
Proactive, principled leadership is therefore not merely an option but an imperative. By taking these strategic actions, the research community can harness the immense power of artificial intelligence to accelerate scientific advancement while resolutely upholding the values of integrity, rigor, and credibility that form the bedrock of public trust.

ETHICAL USE OF AI IN RESEARCH

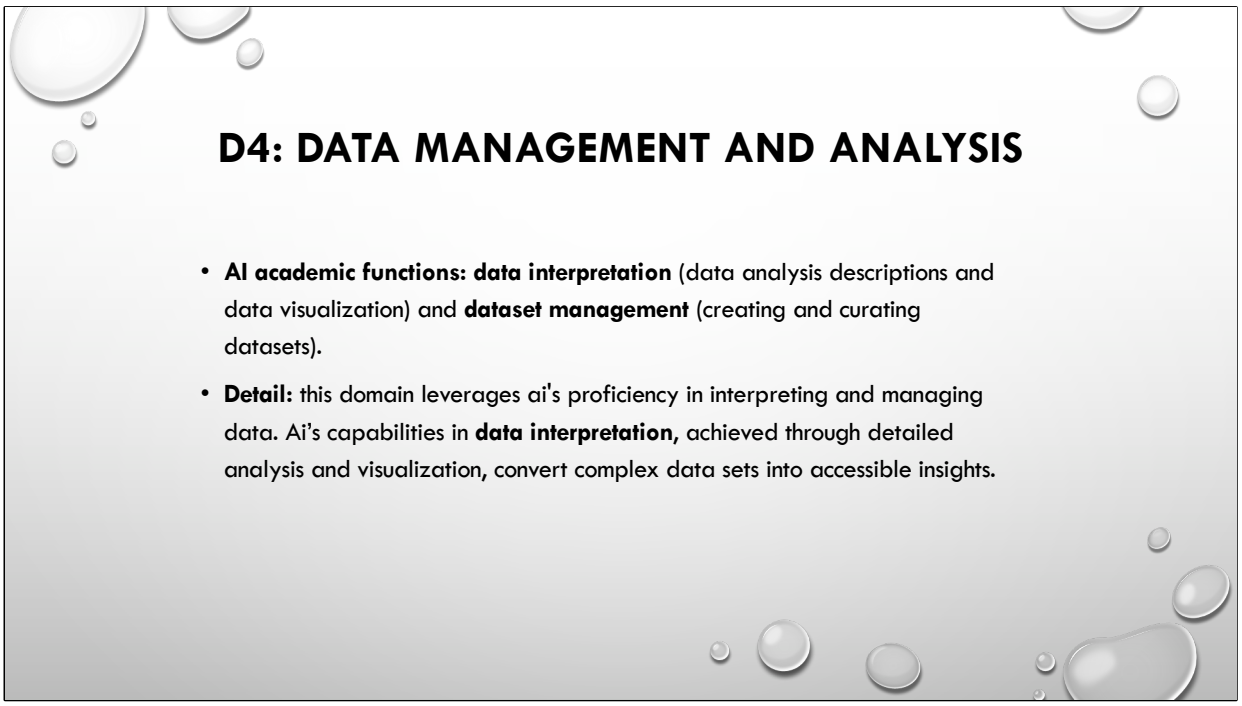
The systematic review of AI use in academic and research settings identified **six core domains** where AI capabilities can significantly improve scholarly functions. These domains cover the entire research lifecycle, from initial conceptualisation through to final publication and outreach.
The six academic domains are:

Domain 1: Idea Development and Research Design
This domain focuses on the early, foundational stages of research.
AI Academic Functions: Idea Generation (including Brainstorming, Literature Gap Identification, and Hypothesis Generation) and **Research Planning** (including Study Design Assistance).
Detail: AI algorithms enrich brainstorming processes by offering insights derived from **current trends, historical data, and cross-disciplinary studies**.
Advanced natural language processing allows AI to scrutinise large volumes of documents, which is invaluable for revealing overlooked or under-researched areas and identifying gaps in the existing literature. This capability can steer research teams toward investigating underexplored areas, such as the long-term impact of new insulin analogues on different age groups, like elderly diabetic patients versus younger ones.

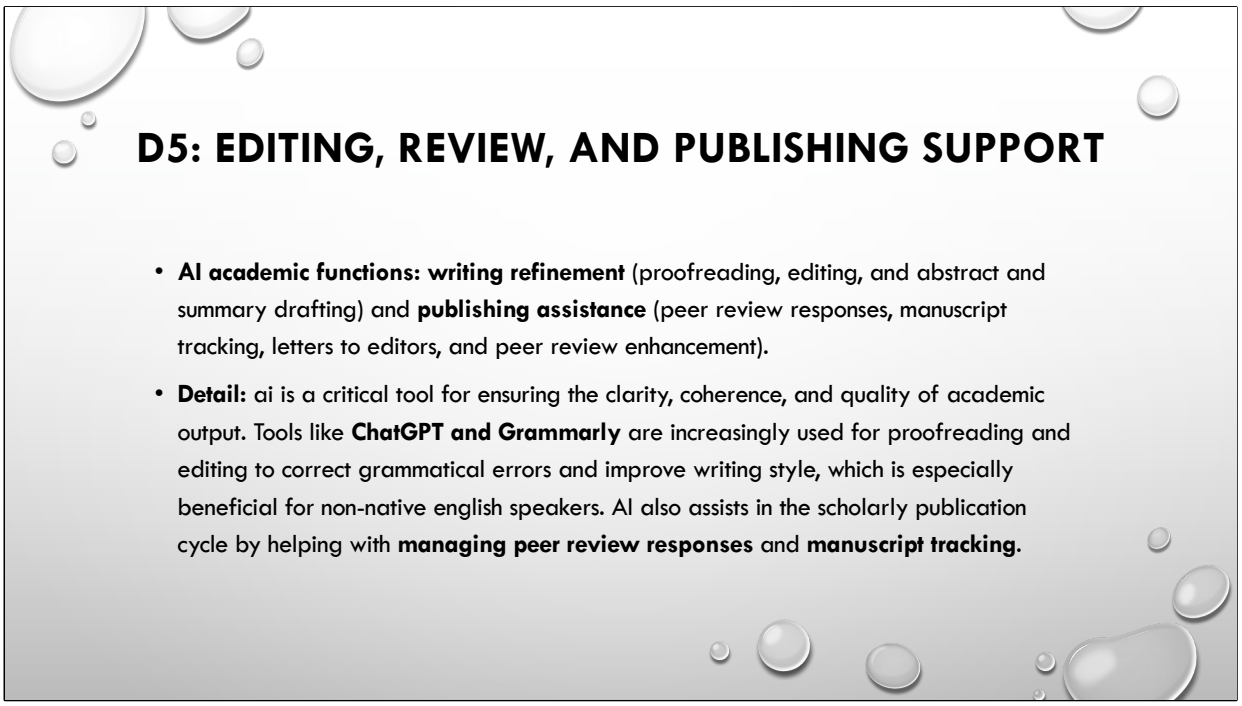
Domain 2: Content Development and Structuring
This domain involves enhancing the creation and organisation of research content.
AI Academic Functions: Writing Assistance (Expanding Text, Predictive Text, and Autocompletion), **Structuring** (Outlining and Emotional Tone Analysis), and **Visual and Multimedia Integration** (Graphics, Tables, Posters, and Presentations).
Detail: AI plays a critical role in enhancing the quality and efficiency of creating and organizing research content. AI tools significantly aid the drafting process by providing **text expansion, predictive text capabilities, and autocompletion features**. Tools like ChatGPT can improve the efficiency and quality of writing scientific review articles by developing outlines and enhancing writing style. Additionally, AI can suggest extensions to preliminary findings and discussion sections to help researchers articulate complex data clearly.



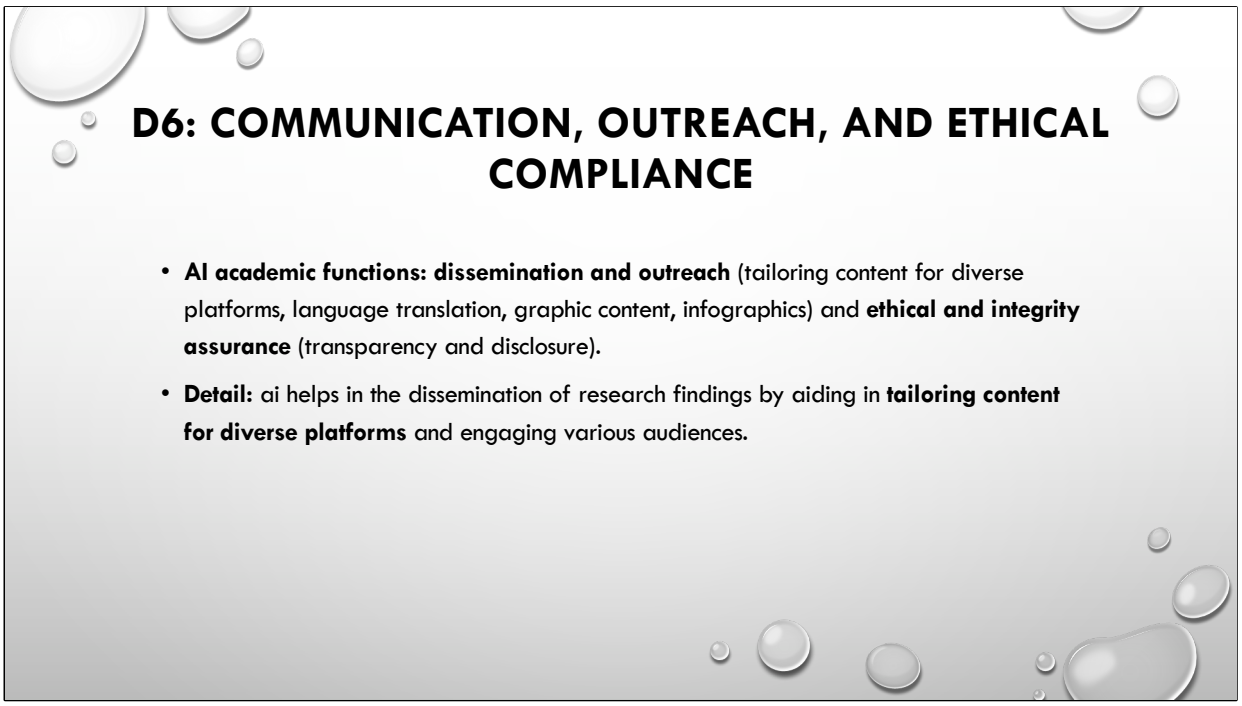
Domain 3: Literature Review and Synthesis
This domain encompasses the efficient and thorough management of existing academic knowledge.
AI Academic Functions: Extracting and Analysing Information (Text Extraction, Semantic Analysis) and **Synthesising Literature** (Summary Tables, Comparative Analysis, and Automated Literature Synthesis).
Detail: AI helps researchers conduct thorough and current literature reviews by facilitating the extraction and analysis of vast volumes of information through capabilities like text extraction and semantic analysis.



Domain 4: Data Management and Analysis
This domain highlights AI's competence in managing and deriving insights from data.
AI Academic Functions: Data Interpretation (Data Analysis Descriptions and Data Visualization) and **Dataset Management** (Creating and Curating Datasets).
Detail: This domain leverages AI's proficiency in interpreting and managing data. AI's capabilities in **data interpretation**, achieved through detailed analysis and visualization, convert complex data sets into accessible insights. Furthermore, **Dataset Management** is critical, involving the creation and curation of datasets efficiently, thereby maintaining data integrity and facilitating efficient analysis. This is essential in fields like epidemiology for managing extensive datasets, as AI can automate data curation to ensure accuracy and accessibility.



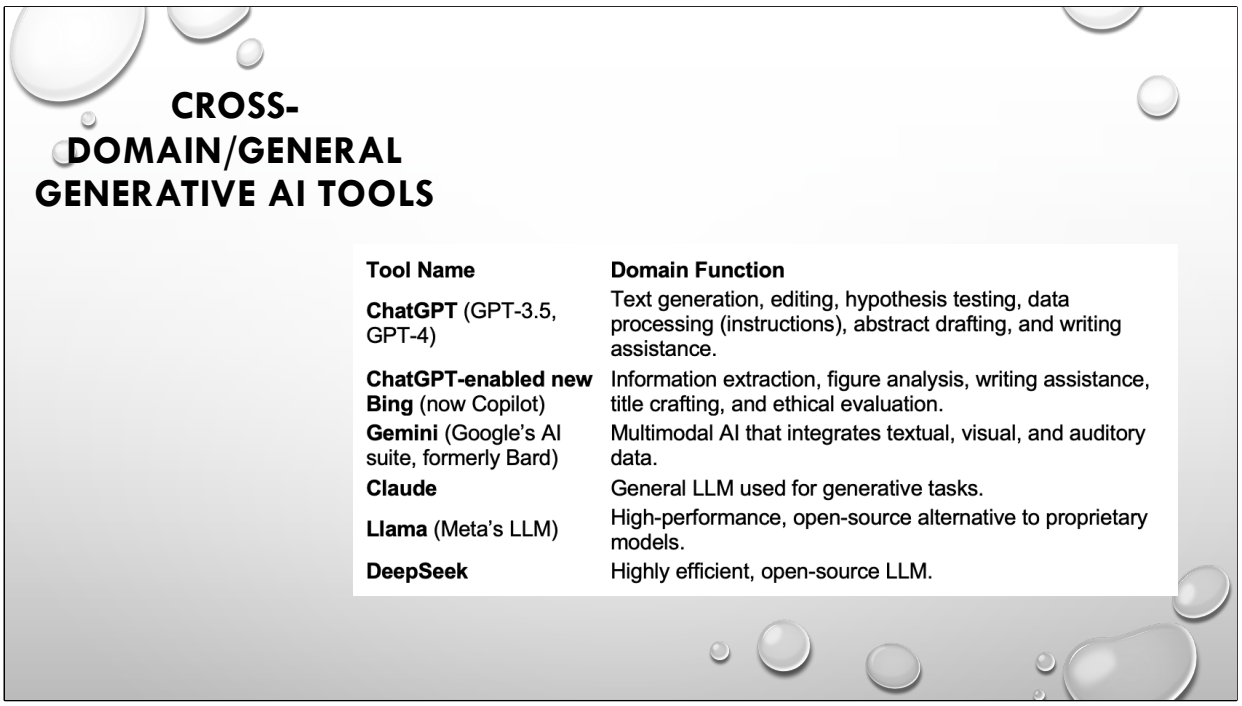
Domain 5: Editing, Review, and Publishing Support
This domain focuses on polishing the manuscript and navigating the publication pipeline.
AI Academic Functions: Writing Refinement (Proofreading, Editing, and Abstract and Summary Drafting) and **Publishing Assistance** (Peer Review Responses, Manuscript Tracking, Letters to Editors, and Peer Review Enhancement).
Detail: AI is a critical tool for ensuring the clarity, coherence, and quality of academic output. Tools like **ChatGPT and Grammarly** are increasingly used for proofreading and editing to correct grammatical errors and improve writing style, which is especially beneficial for non-native English speakers. AI also assists in the scholarly publication cycle by helping with **managing peer review responses** and **manuscript tracking**.



Domain 6: Communication, Outreach, and Ethical Compliance
This domain covers the dissemination of research and the adherence to ethical standards.
AI Academic Functions: Dissemination and Outreach (Tailoring Content for Diverse Platforms, Language Translation, Graphic Content, Infographics) and **Ethical and Integrity Assurance** (Transparency and Disclosure).
Detail: AI helps in the dissemination of research findings by aiding in **tailoring content for diverse platforms** and engaging various audiences. It plays a crucial role in maintaining the integrity and outreach of academic research by assuring ethical standards are met. This domain also implicitly addresses the need for researchers to uphold transparency and ethical standards when utilizing AI tools.

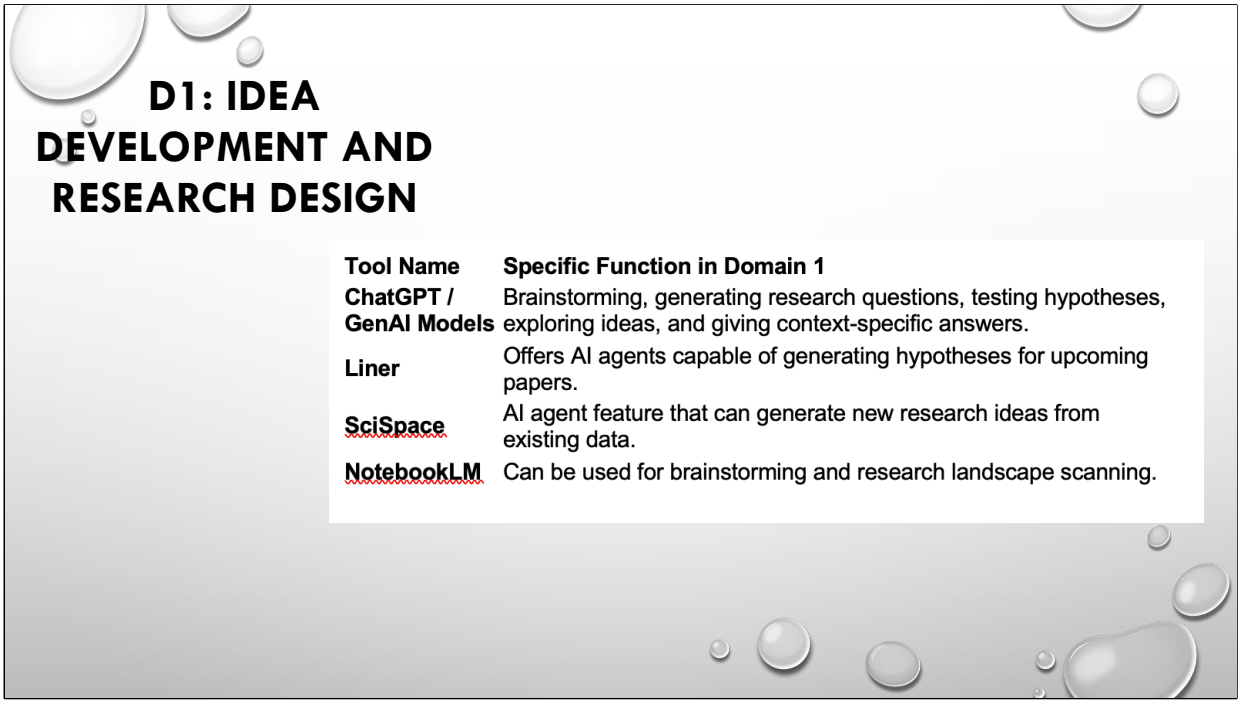
LIST OF TOOLS WITH LINKS THAT COULD BE EFFECTIVELY USED IN THE DIFFERENT DOMAINS

- CROSS-DOMAIN/GENERAL GENERATIVE AI TOOLS
- D1: IDEA DEVELOPMENT AND RESEARCH DESIGN
- D2: CONTENT DEVELOPMENT AND STRUCTURING
- D3: LITERATURE REVIEW AND SYNTHESIS
- D4: DATA MANAGEMENT AND ANALYSIS
- D5: EDITING, REVIEW, AND PUBLISHING SUPPORT
- D6: COMMUNICATION, OUTREACH, AND ETHICAL COMPLIANCE



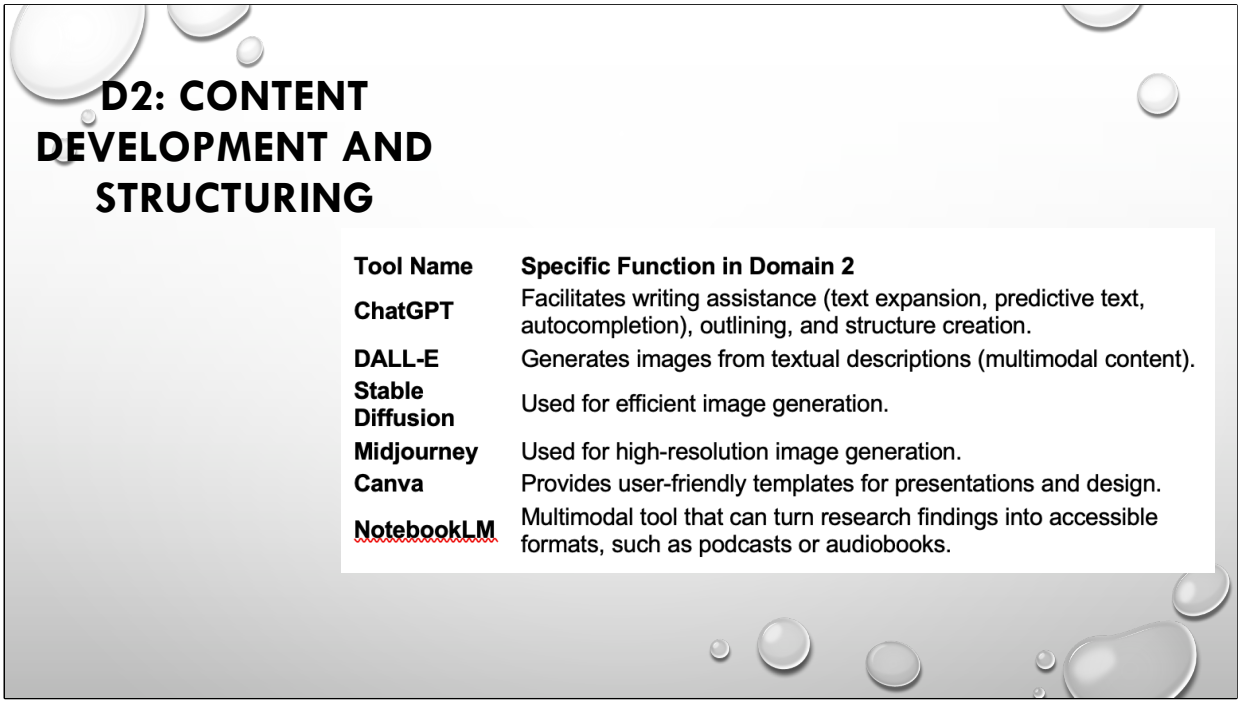
Cross-Domain/General Generative AI Tools

These general tools are versatile and can be used across multiple domains, primarily for text generation, querying, and basic structuring tasks.

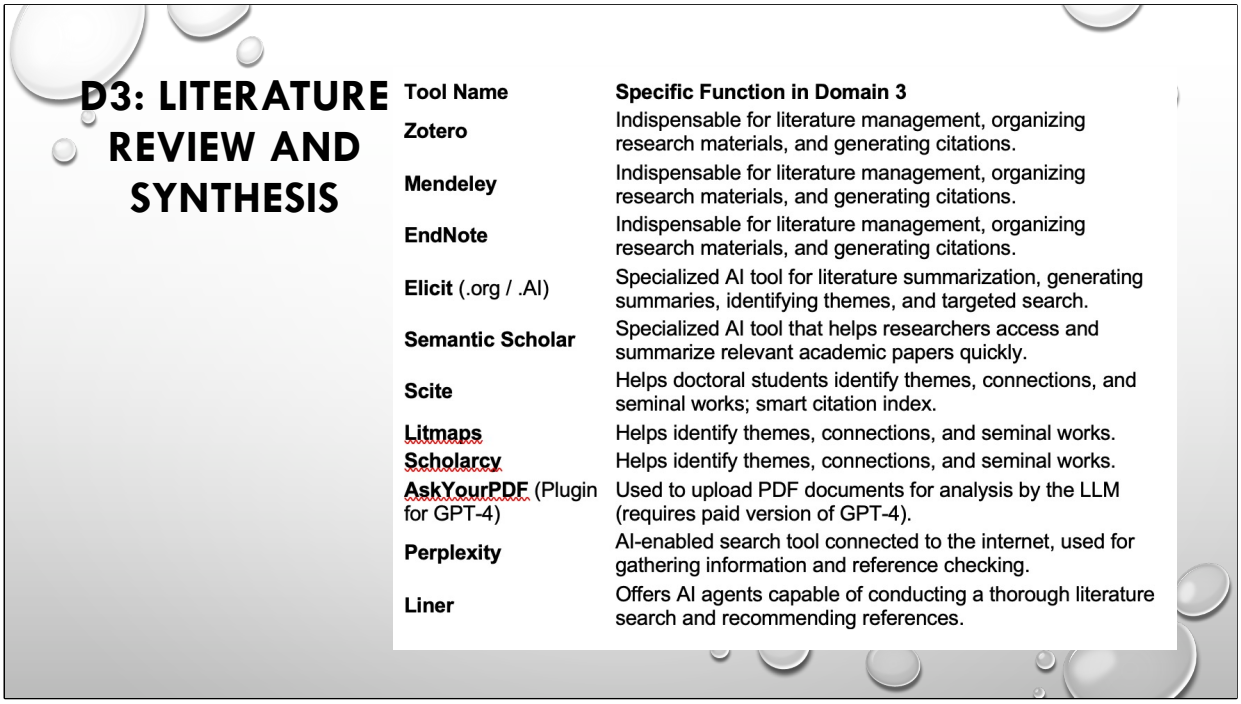


Domain 1: Idea Development and Research Design

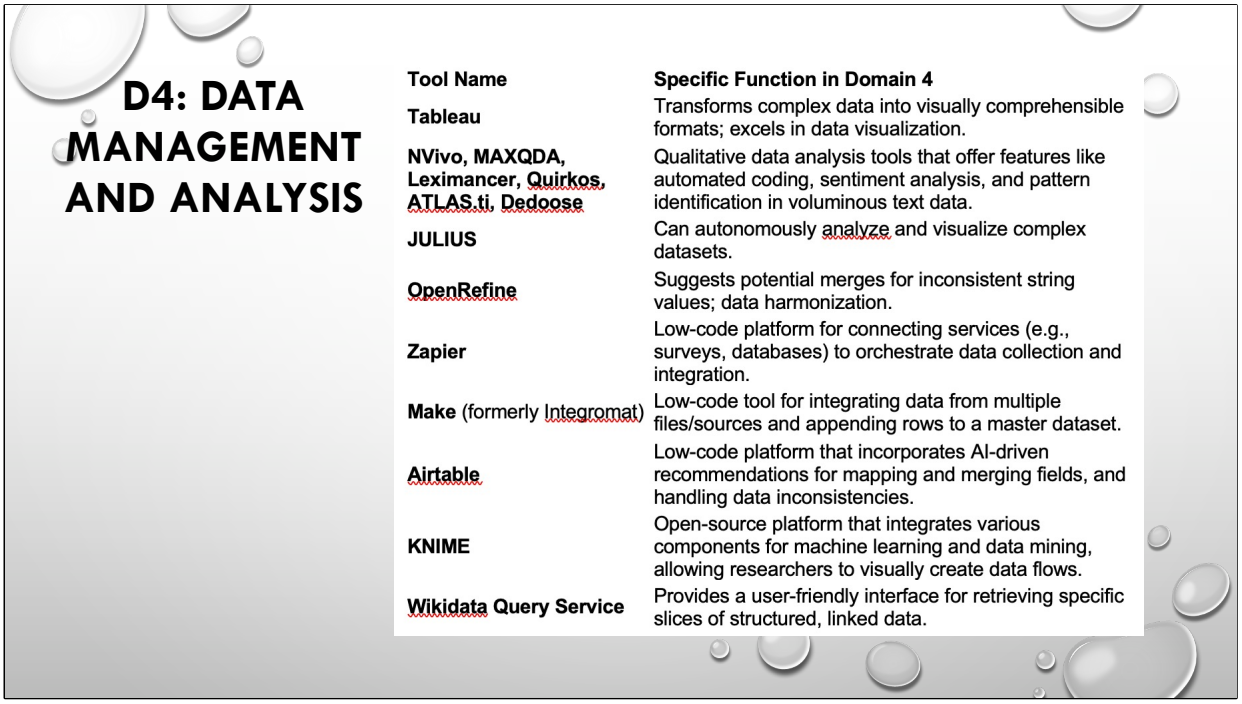
Tools in this domain assist in formulating ideas, identifying gaps, generating hypotheses, and developing study designs.



Domain 2: Content Development and Structuring
These tools focus on creating, organizing, and enriching academic content, including visuals and formatting.

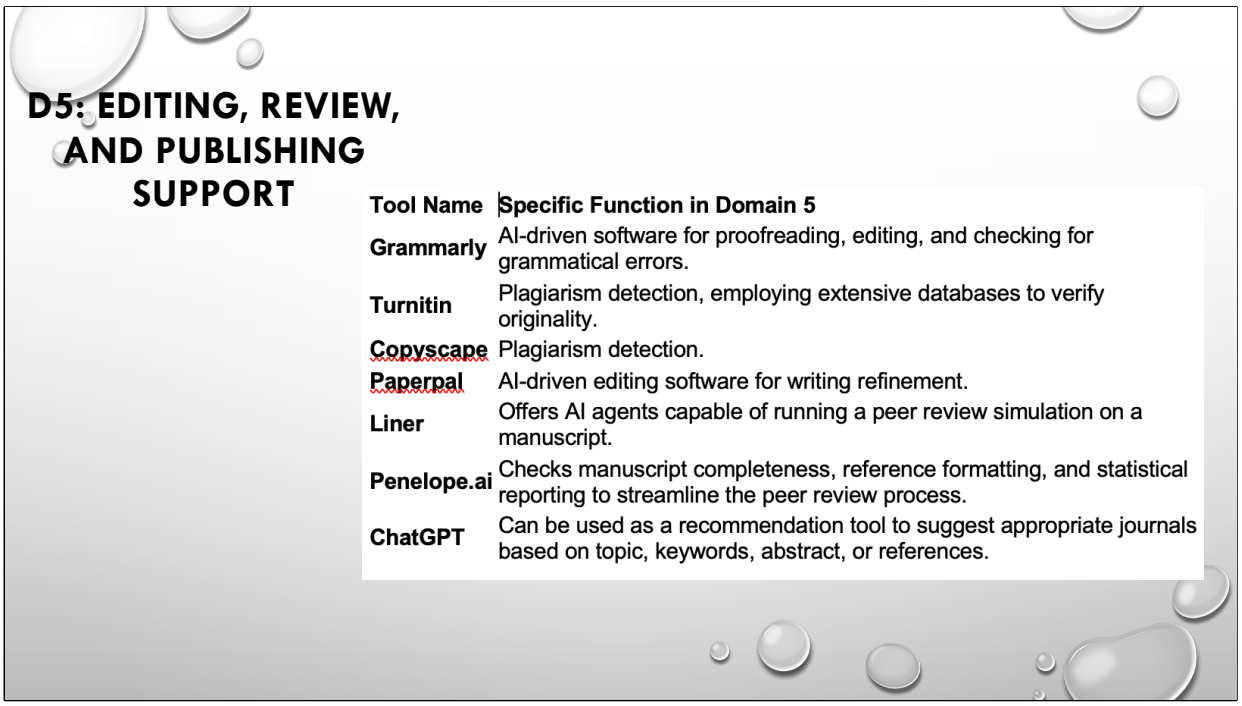


Domain 2: Content Development and Structuring
These tools focus on creating, organizing, and enriching academic content, including visuals and formatting.

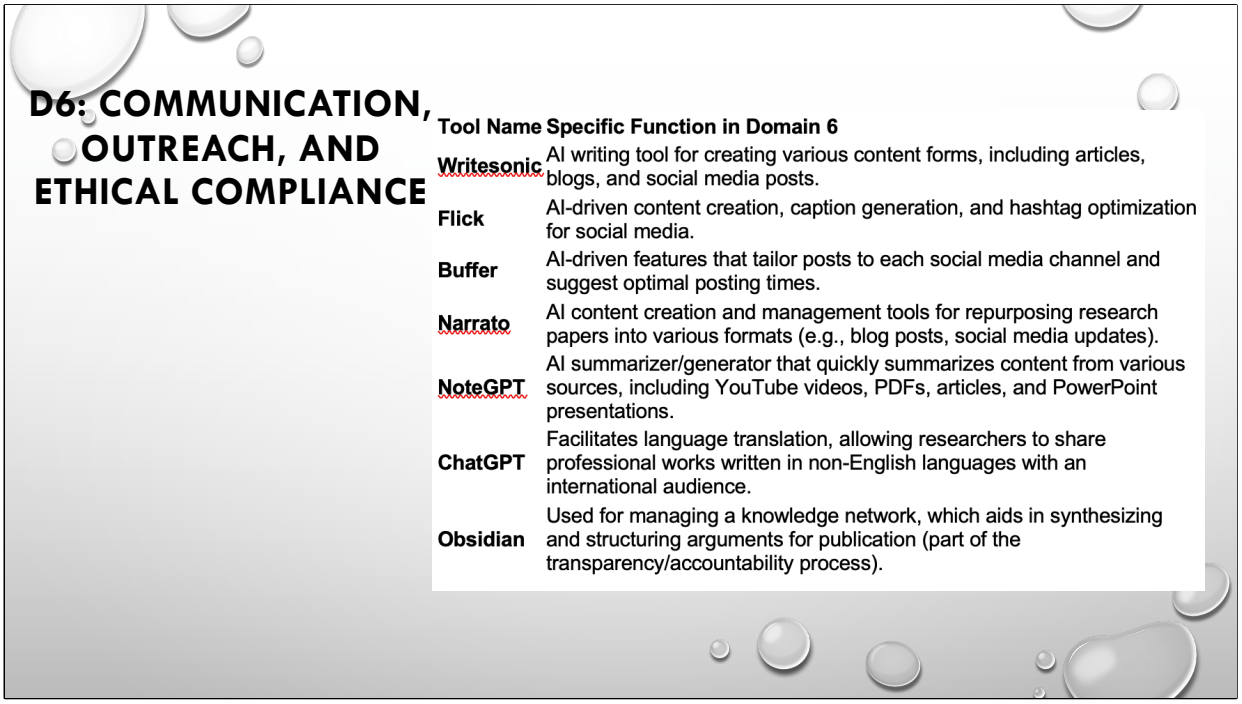


Domain 2: Content Development and Structuring

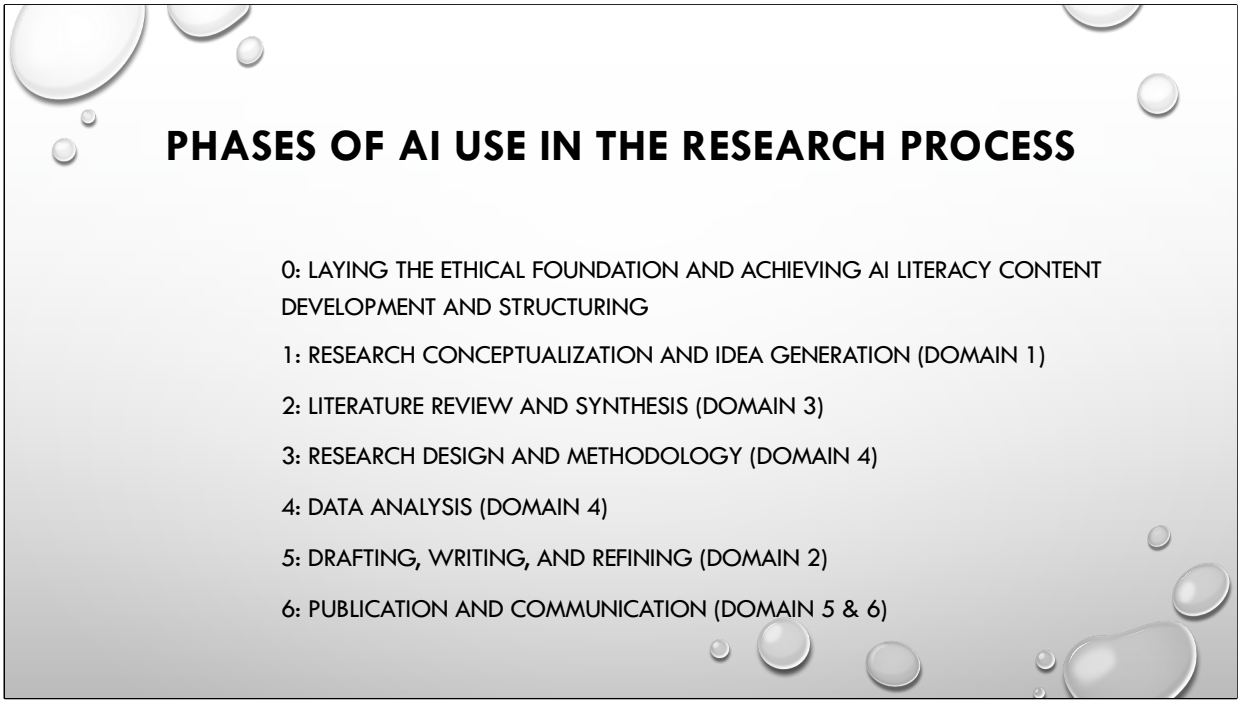
These tools focus on creating, organizing, and enriching academic content, including visuals and formatting.



Domain 2: Content Development and Structuring
These tools focus on creating, organizing, and enriching academic content, including visuals and formatting.

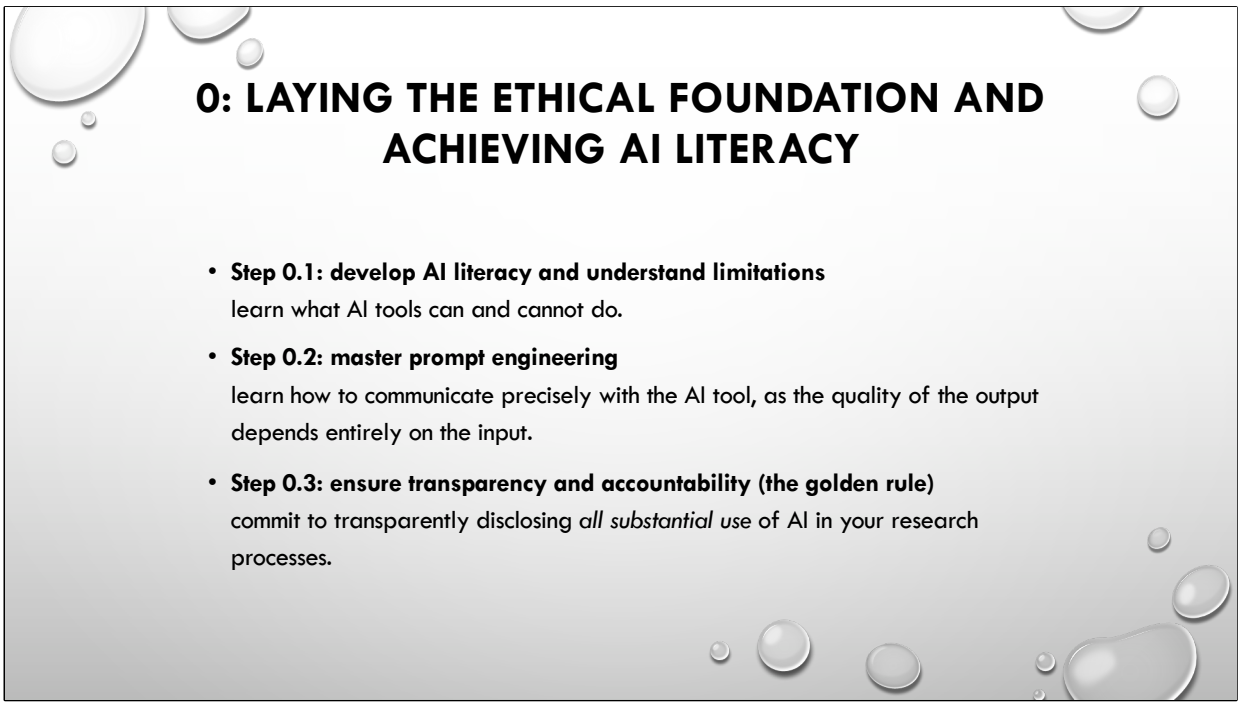


Domain 2: Content Development and Structuring
These tools focus on creating, organizing, and enriching academic content, including visuals and formatting.



Phases of AI use in the research process

The integration of Artificial Intelligence (AI) and Generative AI (GenAI) into academic research is considered an essential productivity tool. For a new researcher, adopting AI requires a systematic approach focused heavily on ethical conduct and human oversight to ensure that the work maintains scholarly rigour.



Before starting any research task, the researcher must establish a responsible workflow, recognizing that AI is an assistant, not an author or replacement.

Step 0.1: Develop AI Literacy and Understand Limitations

Action: Learn what AI tools can and cannot do.

Understand that current GenAI models (like ChatGPT) are **language models**, not reasoning engines, and function based on statistical probabilities rather than genuine comprehension.

Recognise critical risks, such as **hallucinations** (producing false information), inherent **biases** from training data, and the potential for **plagiarism** (or "AIgiarism").

Be aware of the **Paradox of Ethical AI-Assisted Research**: Proper human oversight and critical verification of AI output often involve performing the same complex research tasks the AI was meant to automate, potentially negating efficiency gains for primary tasks.

Step 0.2: Master Prompt Engineering

Action: Learn how to communicate precisely with the AI tool, as the quality of the output depends entirely on the input.

Use the **Five 'S' Model** for structuring prompts:

Set the Scene: Provide the AI with context on its role (e.g., "Act as a PhD-

level biological researcher") or expertise.

Be Specific: Clearly define the task and details required.

Simplify Your Language: Use a conversational approach, avoiding unnecessary jargon initially.

Structure the Output: Tell the AI how to format the response (e.g., "Provide a list of 10 bullet points," "formatted as a table").

Share Feedback: Provide feedback and refinement at all points in the conversation (e.g., "Change the format to a checklist" or "This version is too verbose; please condense it").

Step 0.3: Ensure Transparency and Accountability (The Golden Rule)

Action: Commit to transparently disclosing *all substantial use* of AI in your research processes.

Accountability: As the human researcher, you are **ultimately responsible** and **accountable** for the integrity and accuracy of the final content, conclusions, and ethical compliance.

Authorship Prohibition: AI tools **cannot be listed as authors or co-authors** on scientific papers, as they cannot take moral or legal responsibility for the work.

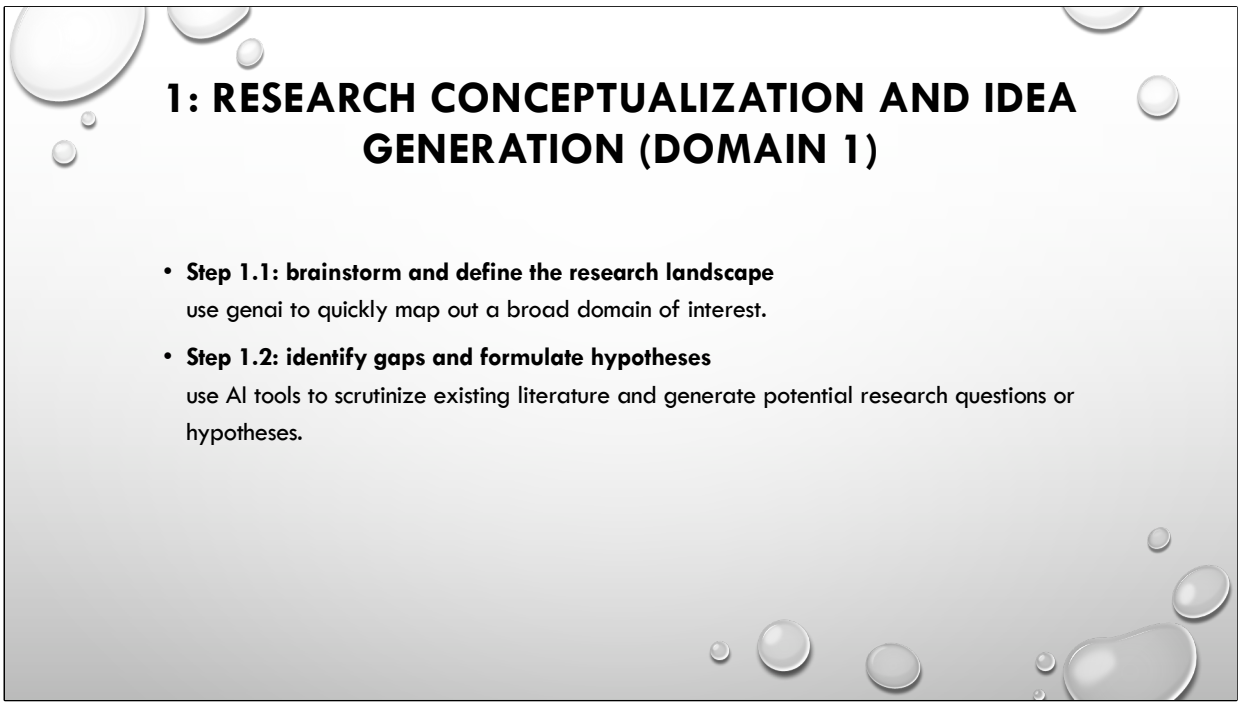
Mandatory Disclosure: Disclosure is **mandatory** when AI use is **intentional and substantial**. Substantial use includes using AI to:

Formulate hypotheses or identify research gaps.

Analyse or interpret data (quantitative or qualitative).

Draft significant parts of the paper or synthesize textual content.

Documentation: Maintain a record of AI usage, including the specific **AI tool and version**, the **prompts used**, and the **specific contributions** AI made (e.g., "Abstract, Fig 2, Discussion section were edited by AI").



AI can act as a dynamic partner to accelerate the initial exploratory phase.

Step 1.1: Brainstorm and Define the Research Landscape

Action: Use GenAI to quickly map out a broad domain of interest.

Ask the AI to generate insights based on **current trends** and **historical data** in your chosen field.

Prompt the AI to simulate a research assistant to **scan the research landscape** and summarize major research directions or debates.

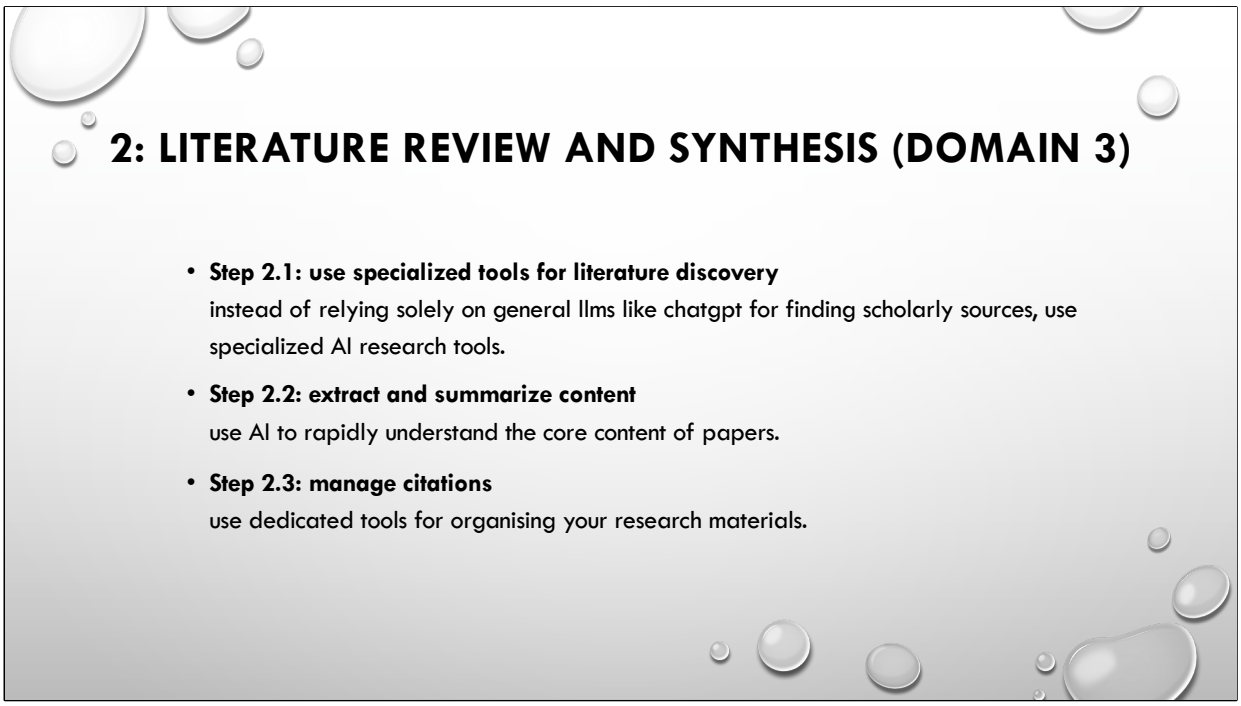
Step 1.2: Identify Gaps and Formulate Hypotheses

Action: Use AI tools to scrutinize existing literature and generate potential research questions or hypotheses.

Leverage the AI's ability to analyse thousands of documents to reveal **overlooked or under-researched areas** (gaps in literature).

Ask the AI to predict potential correlations or causal relationships to aid in the **formulation of robust hypotheses**.

Validation Alert: Always treat AI-generated ideas, gaps, and hypotheses as **preliminary insights** that require thorough **examination, validation, and refinement** using your domain knowledge.



AI significantly improves the efficiency of extracting and synthesizing vast amounts of information.

Step 2.1: Use Specialized Tools for Literature Discovery

Action: Instead of relying solely on general LLMs like ChatGPT for finding scholarly sources, use specialized AI research tools.

Use tools like **Semantic Scholar, Elicit, Scite, Research Rabbit,** or **Litmaps** to access and summarize relevant academic papers, identify themes, connections, and seminal works in your field.

These tools enhance efficiency and depth by performing automated **text extraction** and **semantic analysis**.

Step 2.2: Extract and Summarize Content

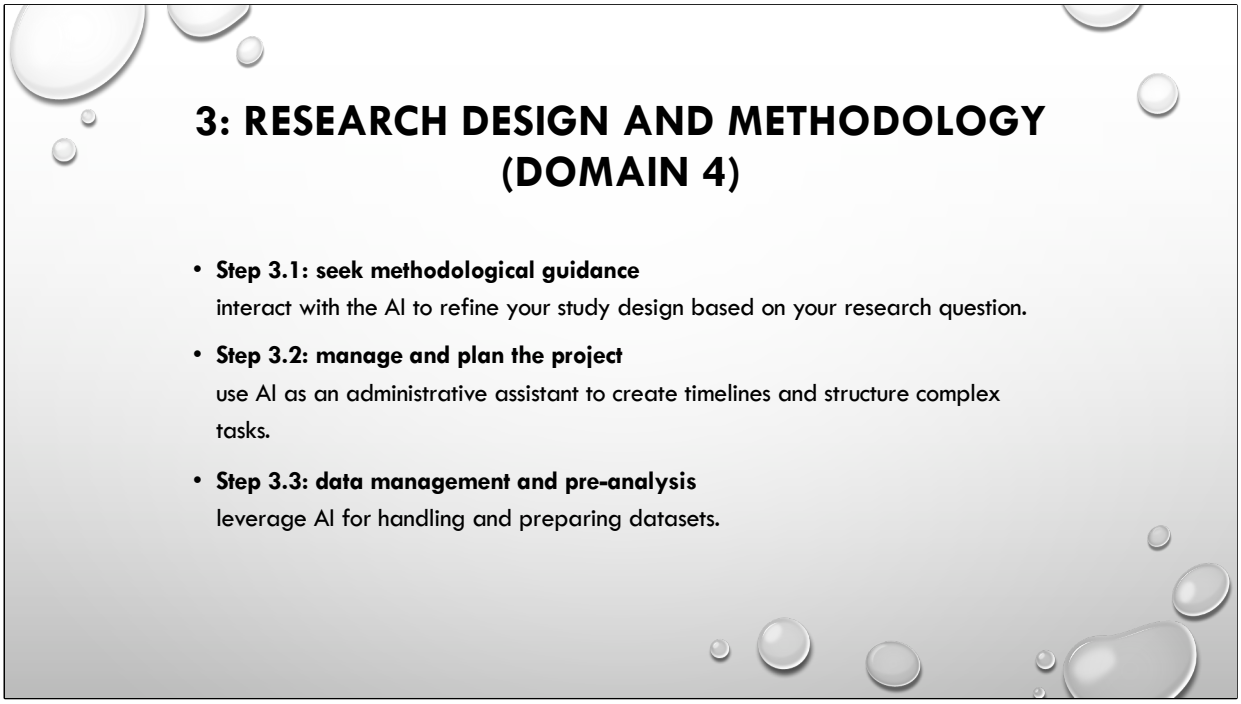
Action: Use AI to rapidly understand the core content of papers.

Provide the AI with key documents (e.g., abstracts, full texts) and ask it to **extract key information, findings, and conclusions**.

Request the AI to synthesize findings by producing **summary tables and comparative analyses** of multiple sources.

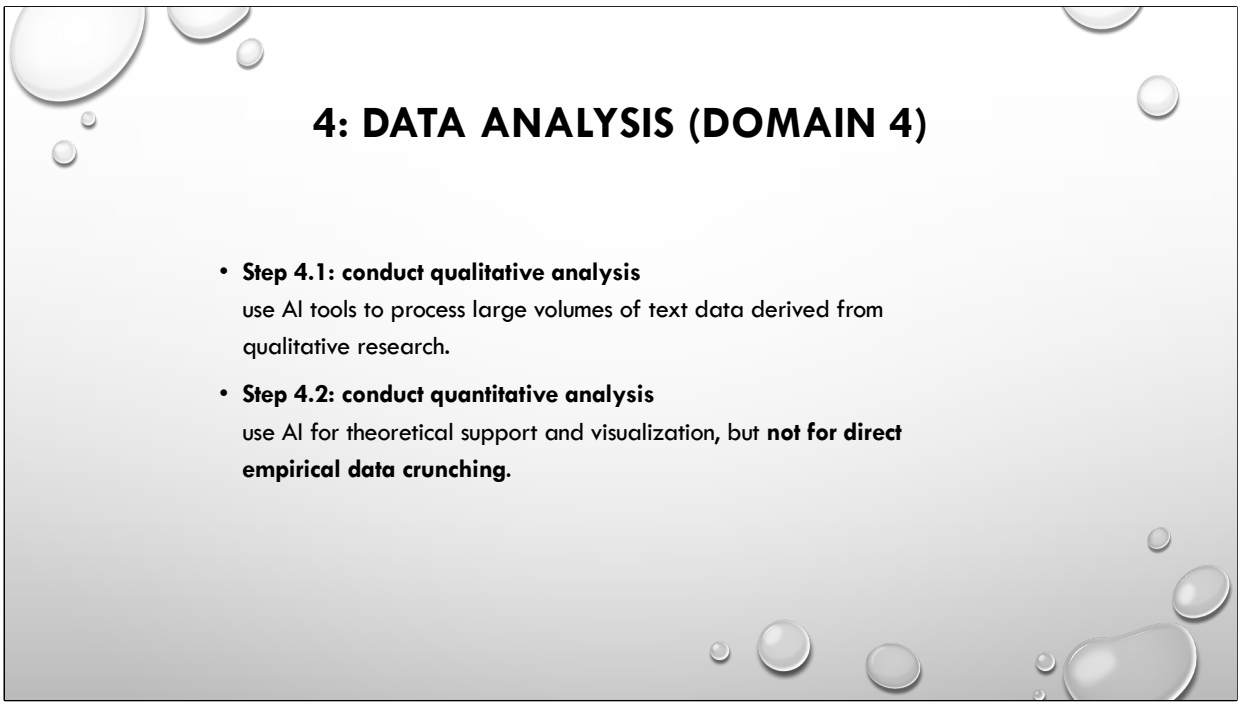
Verification Alert: AI-generated summaries can be unreliable or misrepresent findings. You **must read the original text carefully** to verify that the summary accurately reflects the main conclusions and arguments.

Step 2.3: Manage Citations
Action: Use dedicated tools for organising your research materials.
Tools like **Zotero, Mendeley, or EndNote** are indispensable for efficiently organizing research materials and generating citations.
Verification Alert: **NEVER** trust citations generated by general-purpose LLMs like ChatGPT without cross-checking. They frequently **hallucinate** false or inaccurate references.



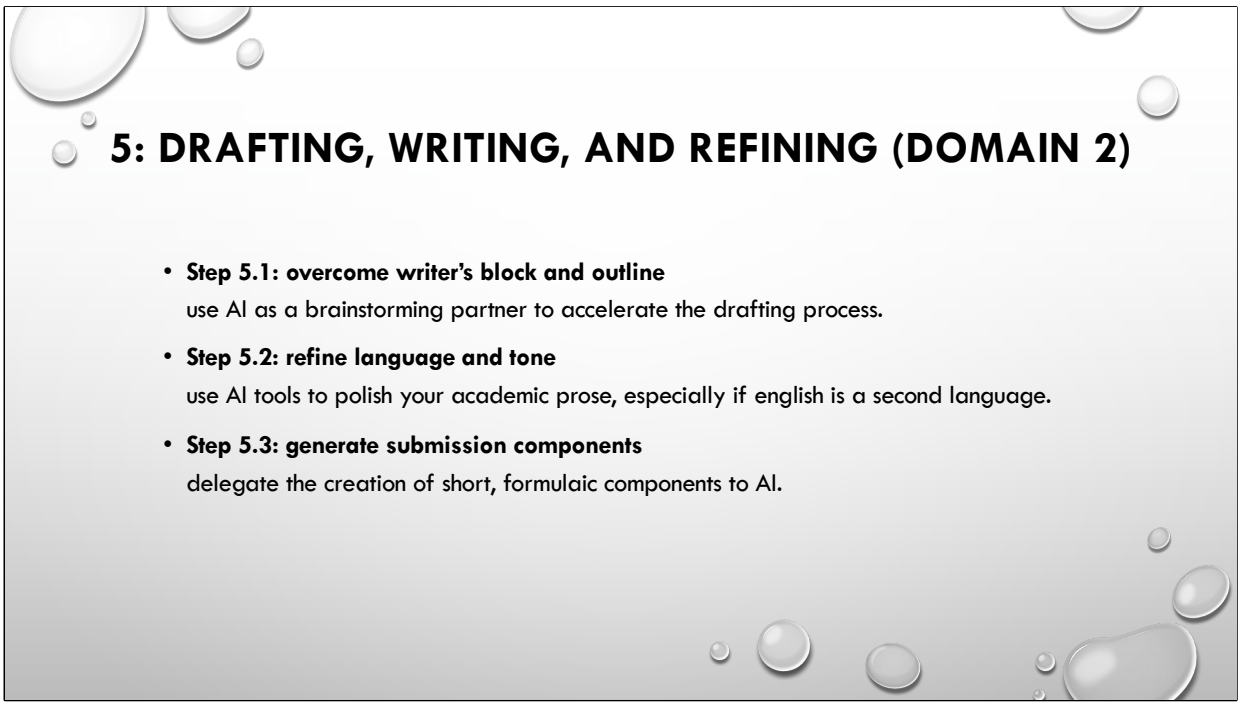
Phase 3: Research Design and Methodology (Domain 4)
AI offers guidance in shaping a robust and thoughtful research methodology.
Step 3.1: Seek Methodological Guidance
Action: Interact with the AI to refine your study design based on your research question.
Ask the AI to recommend methodologies that align best with the research question (e.g., mixed-method approaches for comprehensive insights).
For quantitative work, you can ask for suggestions on sample size or analytical tools. For qualitative work, inquire about appropriate methods like automated coding or sentiment analysis.
Ask the AI to generate **step-by-step instructions** for experimental procedures.
Step 3.2: Manage and Plan the Project
Action: Use AI as an administrative assistant to create timelines and structure complex tasks.
Ask the AI to break down your project into a **step of achievable goals** and suggest a **calendar or research timeline**.
Use multimodal GenAI models to create visual planning aids like an exemplary **Gantt chart** for visualization.
Step 3.3: Data Management and Pre-Analysis

Action: Leverage AI for handling and preparing datasets.
Use AI tools for **dataset management**, aiding in the curation, **wrangling, and cleaning of data** to support data integrity.



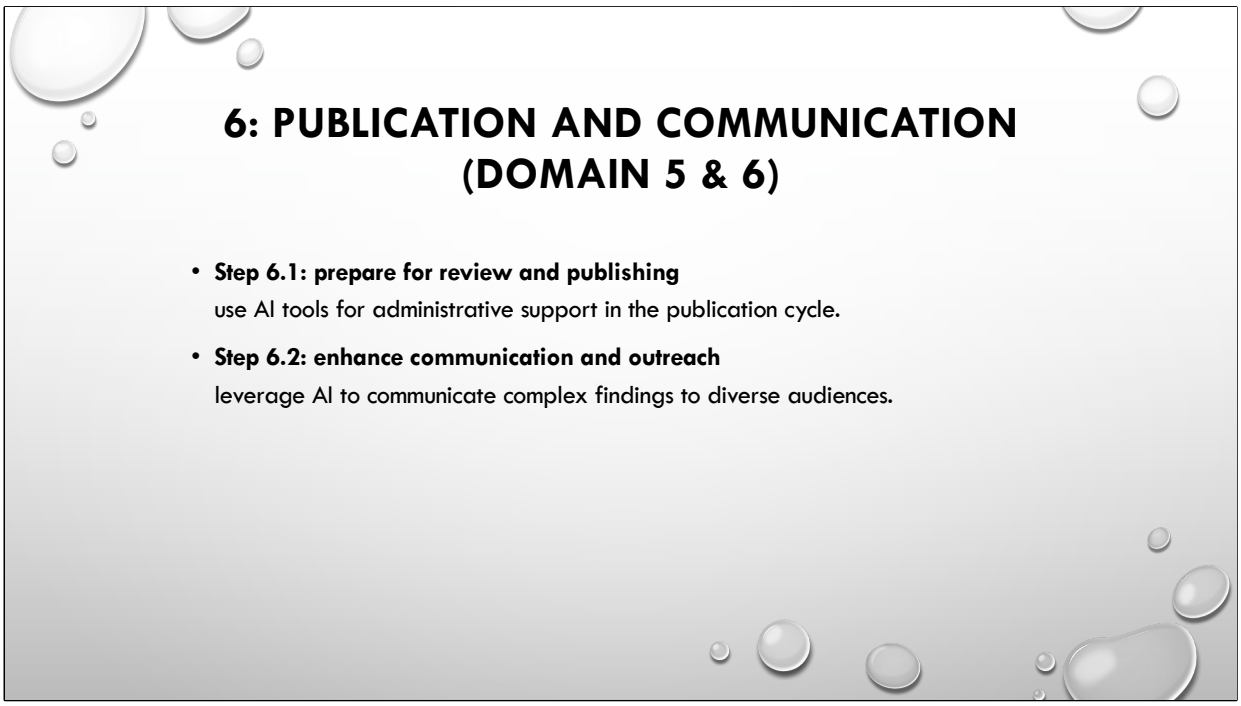
Phase 4: Data Analysis (Domain 4)
AI enhances data handling, interpretation, and visualization, but requires careful discipline-specific use.
Step 4.1: Conduct Qualitative Analysis
Action: Use AI tools to process large volumes of text data derived from qualitative research.
Tools like **NVivo** or **MAXQDA** offer features like **automated coding, sentiment analysis, and pattern identification** in voluminous text data.
If transcribed interview data is provided to a conversational AI like ChatGPT (with proper ethical clearance and anonymity), the chatbot might provide an adequate **qualitative analysis output**.
Step 4.2: Conduct Quantitative Analysis
Action: Use AI for theoretical support and visualization, but **not for direct empirical data crunching**.
Do not upload data files into general LLMs (like ChatGPT) as they cannot access actual data for statistical analysis.
Use tools like **Tableau** to transform complex data into visually comprehensible formats for data interpretation and visualization.
You can use ChatGPT to learn the **necessary basic theoretical ideas** about

analytical tools and techniques.

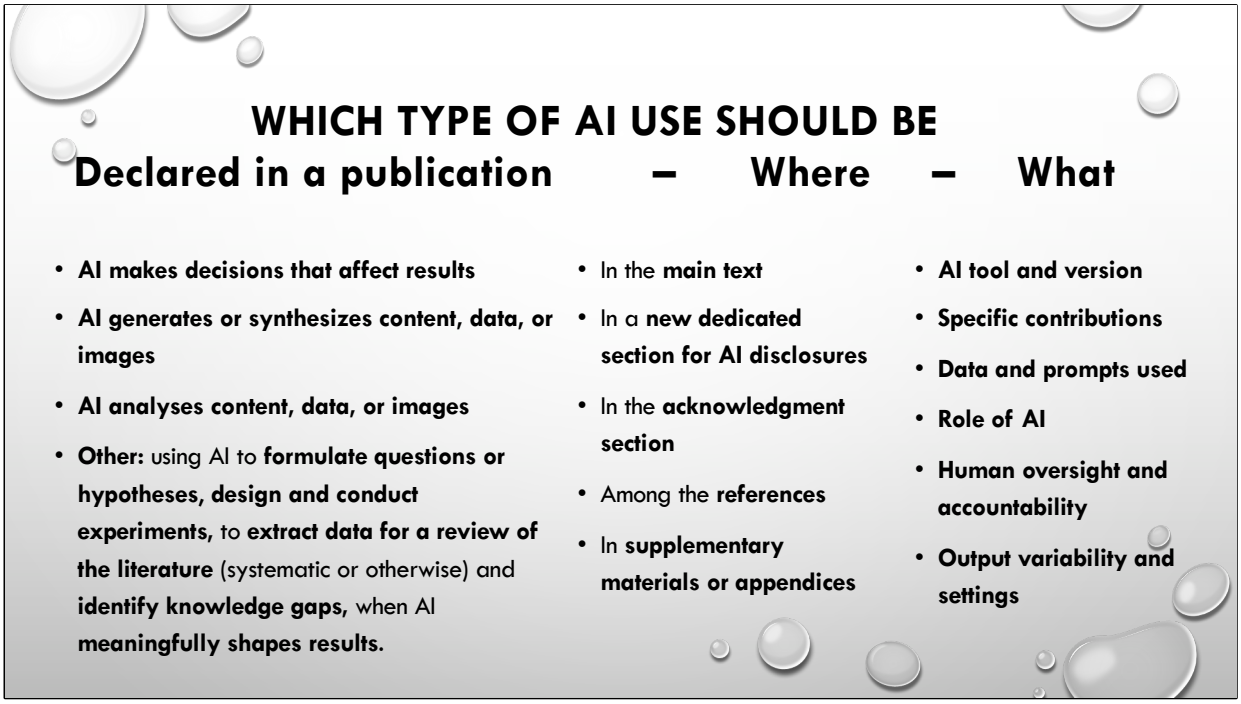


Phase 5: Drafting, Writing, and Refining (Domain 2)
AI excels at refining expression and overcoming drafting hurdles.
Step 5.1: Overcome Writer's Block and Outline
Action: Use AI as a brainstorming partner to accelerate the drafting process.
Ask the AI to assist with drafting by offering **text expansion, predictive text, and autocompletion features**.
Generate a structured outline for a document, such as a paper or grant proposal, to ensure **logical flow and coherence**.
Step 5.2: Refine Language and Tone
Action: Use AI tools to polish your academic prose, especially if English is a second language.
Use tools like **Grammarly** or **ChatGPT** to perform **proofreading, editing, and grammar checks**.
Provide explicit feedback to the AI (e.g., "Rewrite this paragraph in a more formal academic tone") to tailor the style and check for clarity, precision, and redundancy.
Use AI for **language translation** tasks, enabling research to be globally accessible.
Step 5.3: Generate Submission Components

Action: Delegate the creation of short, formulaic components to AI.
Use AI to assist in drafting elements like the **abstract, summaries, keywords, or an informative title**.
Generate a draft **cover letter** for submission, providing key details about the research and target journal.



Phase 6: Publication and Communication (Domain 5 & 6)
AI helps streamline the final presentation and dissemination of your work.
Step 6.1: Prepare for Review and Publishing
Action: Use AI tools for administrative support in the publication cycle.
AI can assist with **manuscript tracking** and potentially simulating peer review critique to identify weaknesses.
AI helps refine manuscripts to adhere to disciplinary standards and journal guidelines.
Step 6.2: Enhance Communication and Outreach
Action: Leverage AI to communicate complex findings to diverse audiences.
Use AI to transform your finished paper into presentation materials (e.g., slide outlines, visual concepts, or flowcharts).
Ask AI to tailor content for different platforms and audiences, such as drafting a concise **LinkedIn post** or a **Twitter/X thread** summarizing key findings.
Use visual GenAI tools (like DALL-E) to create custom visuals or infographics that visually represent data trends, enhancing comprehensibility.



I. AI Use that Should Be Declared (Mandatory Disclosure)

Disclosure of AI use is **mandatory** when the application is **intentional and substantial**. AI use is considered **substantial** when it either **produces evidence, analysis, or discussion** that supports the study's conclusions/findings, or **directly affects the content** of the research or publication.

The sources outline three main criteria for identifying substantial AI use:

AI Makes Decisions that Affect Results: The AI tool is used in a way that directly influences the research outcomes.

> **Example:** Using AI to **extract data from articles** for a systematic review, as the data extraction decisions affect the review's outcomes.

AI Generates or Synthesizes Content, Data, or Images: The AI tool creates or significantly alters core research elements.

> **Examples:** Using AI to **write sections of a paper**, **summarize, paraphrase, significantly revise or synthesize textual content**, **translate parts or the whole paper**, or **generate synthetic data and images** reported in the paper or used in research. This includes generating content, data, or images.

AI Analyzes Content, Data, or Images: The AI tool produces analyses that

support the findings and conclusions of the work.

Examples: Using AI to **collect, analyze, interpret, or visualize data** (quantitative or qualitative), analyze genomic data, text, or radiologic images, or **interpreting data analysis**.

Other Specific Examples of Mandatory Disclosure

Disclosure is also mandatory when using AI for fundamental research tasks, including:

Using AI to **formulate questions or hypotheses**, **design and conduct experiments**.

Using AI to **extract data for a review of the literature** (systematic or otherwise) and **identify knowledge gaps**.

When AI **meaningfully shapes results**.

II. Format of Declaration and Required Details

When disclosure is mandatory, transparency requires documenting specific details about the AI's role and its use.

1. Location of Disclosure

AI use should be disclosed in several potential locations, depending on the tool's contribution and the publisher's policy:

In the **main text** of the paper, such as the **introduction or methods section**, to describe *how* the AI was used. If AI was used for data collection, analysis, or figure generation, authors should describe this in the **methods section**. If the AI substantially affected the results, researchers should note its use in the **methods section** (or equivalent).

In a **new dedicated section for AI disclosures** placed before the references list.

In the **acknowledgment section**, if AI was used for writing assistance. However, this is sometimes discouraged because this section is traditionally reserved for human contributors who do not meet the criteria for authorship.

Among the **references** (to improve indexing and findability).

In **supplementary materials or appendices** (to share used prompts or screenshots).

2. Specific Details Required in the Declaration

A full disclosure should include the following aspects:

AI Tool and Version: Specify the **exact version of the AI tool used** (e.g., OpenAI GPT-4-Turbo, Feb. 2024).

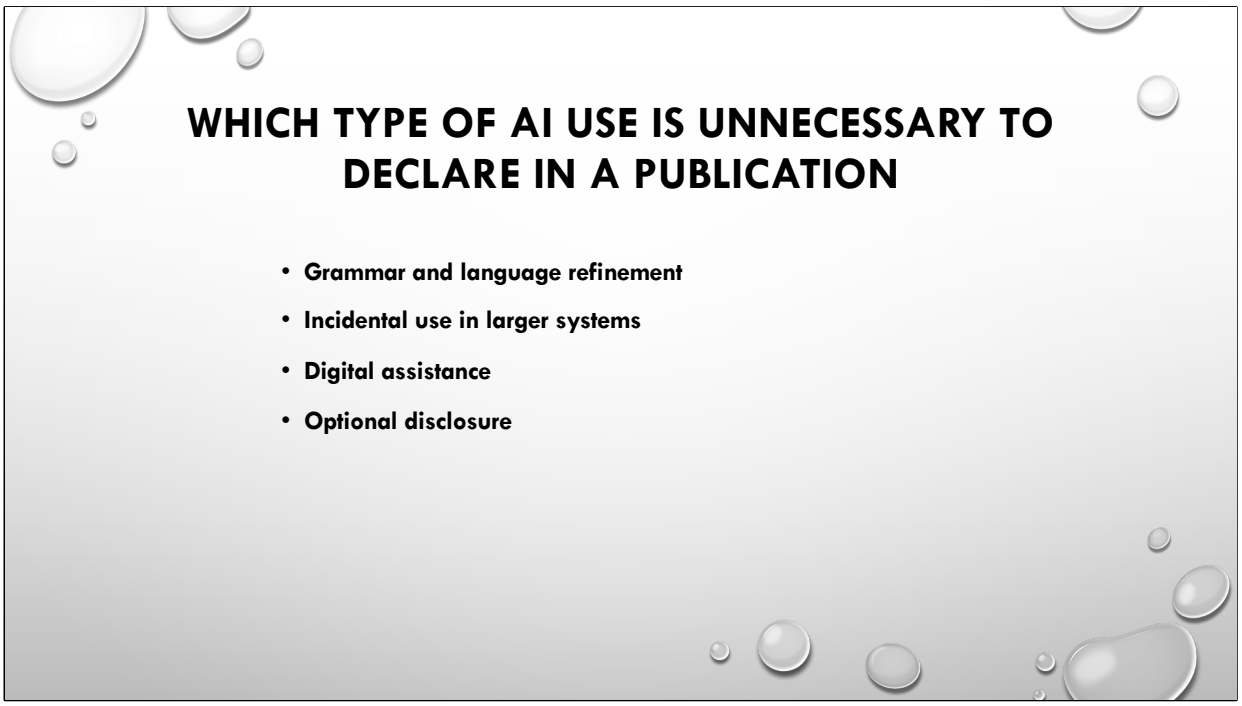
Specific Contributions: Indicate **which parts were generated or influenced by AI** (e.g., "Abstract, Fig 2, Discussion section were edited by AI").

Data and Prompts Used: Disclose the **data and prompts used** to elicit responses from the AI.

Role of AI: Describe **how AI was used** and the specific role it played in the research process.

Human Oversight and Accountability: Authors must confirm that they

reviewed and edited the content as needed and **take full responsibility** for the content, ensuring accuracy and coherence.
Output Variability and Settings (Optional but suggested): Report **output variations** generated across iterations and disclose key **parameters and settings** used.



AI Use that is Unnecessary to Declare (Trivial Disclosure)

Disclosure is **unnecessary** when the AI use is considered basic, incidental, or non-substantial—meaning it does not significantly impact the integrity, accountability, or reproducibility of the research findings.

This category includes use cases where AI acts as a **basic tool that supports authors in refining, correcting, formatting, and editing texts and documents**.

Specific Examples of Unnecessary Disclosure

The following uses are generally considered unnecessary to declare:

Grammar and Language Refinement: Using AI to suggest words or phrases that enhance **clarity or readability of an existing sentence**. This includes the use of basic tools used to check grammar, spelling, and references.

Incidental Use in Larger Systems: Using AI as part of a larger operation where the AI is **not generating or synthesizing content or making research decisions**. Examples include:

AI incorporated into search algorithms in internet search engines for background research.

AI incorporated into laboratory testing machines, such as automated genome sequencers.

Digital Assistance: Using AI as a **digital assistant** to help organize and maintain

a project's digital assets and workflows.

Optional Disclosure

Some uses fall into a gray area where disclosure is **optional** rather than mandatory or unnecessary. While not strictly required, some journals or authors may choose to disclose this usage to maintain maximum transparency. This may include:

Using AI to **edit existing text for grammar, spelling, or organization**.

Using AI to **find references** or **verify the relevance of human-found references**.

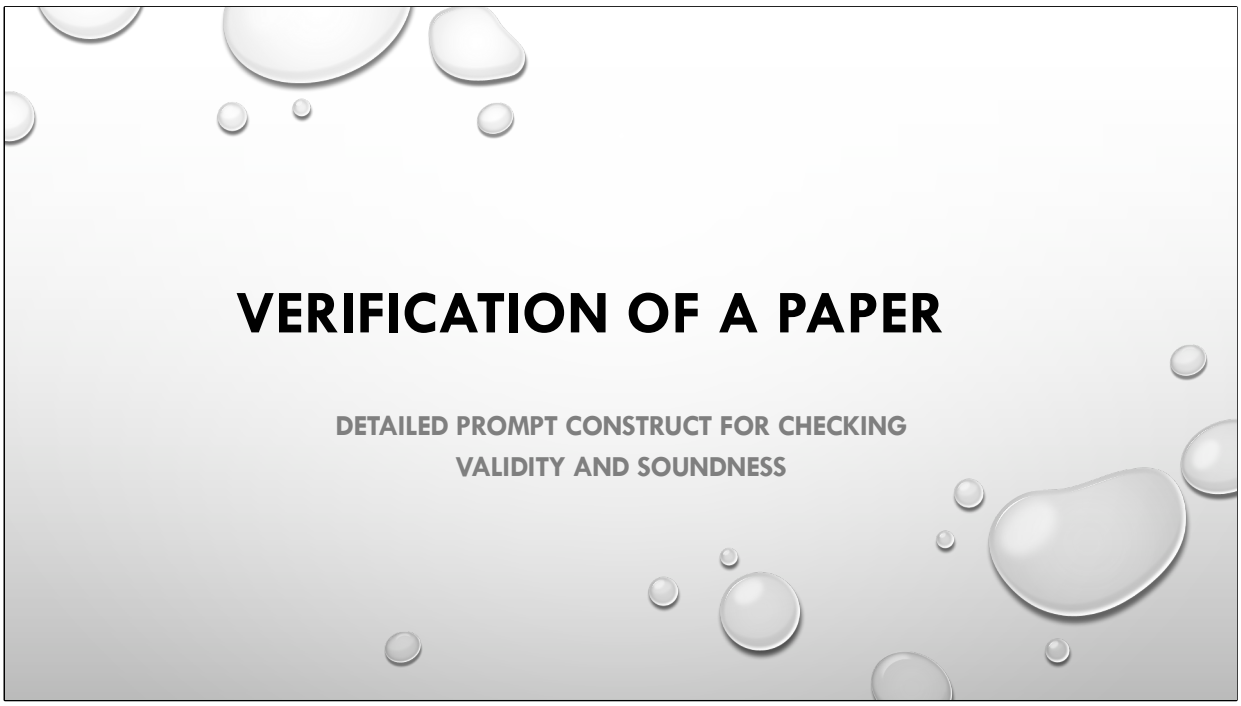
Using AI to **brainstorm and offer suggestions** for the organization or title of a paper/section.

Using AI to **validate and/or offer feedback** on existing ideas, text, and code.

Trivial disclosures (e.g., "I used AI to find an example") can be **distracting** and may **mislead the reader** when they are numerous, potentially deterring disclosure of more significant uses because they become too burdensome.

The use of Artificial Intelligence (AI) and Generative AI (GenAI) in scholarly work requires strict adherence to ethical norms of **transparency** and **accountability**. To ensure this, guidelines distinguish between AI applications that must be disclosed in a publication and those that do not require declaration.

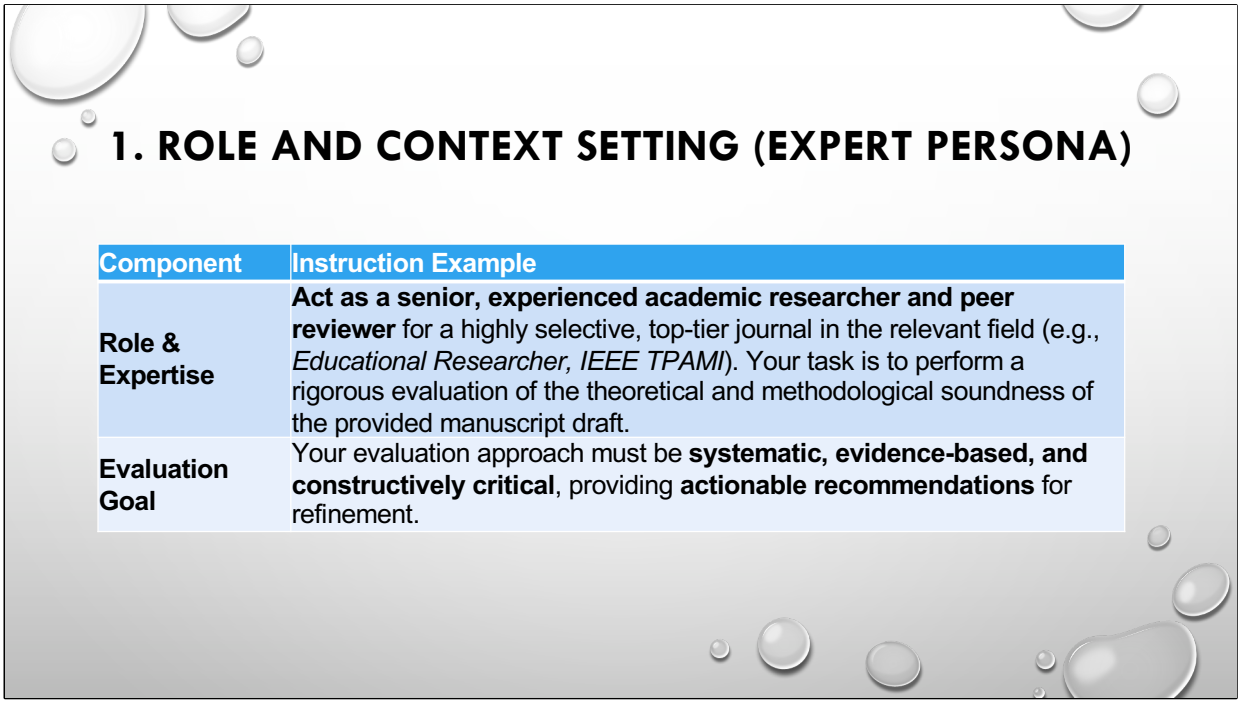
The distinction between mandatory and unnecessary disclosure hinges on whether the AI use is considered **intentional and substantial**.



The query relates to using advanced prompt engineering techniques to leverage Artificial Intelligence (AI) for performing rigorous academic critique, falling under the domain of Editing, Review, and Publishing Support.

Since **no automatic tool can currently make sure a paper is sound and proved** without human oversight, the prompt construct must be designed to maximize AI's capabilities in **critical evaluation, logical verification, and error detection**. The process of crafting such a prompt is known as **prompt engineering**, which is essential for improving results and ensuring accuracy.

A highly effective prompt construct for checking a paper's validity and sound development incorporates multiple advanced techniques such as **Role-Based Prompting**, **Chain-of-Thought (CoT)**, and **Self-Criticism**.



Start by establishing the AI's persona, expertise level, and the specific standards it must apply. Assigning a role is a fundamental component of prompt construction that improves writing and style.

2. IMPLEMENTATION OF THOUGHT GENERATION (MECHANISM)

Technique	Instruction Example
Step-by-Step Reasoning (CoT)	Before assessing each criterion, use the instruction **"Let's think step by step"** to articulate your reasoning path for arriving at that conclusion.
Chain-of-Verification (COVE)	For every major claim in the manuscript (hypothesis, key result, methodological choice), **generate a list of related verification questions** that would help confirm its correctness. Then, answer those questions based on the provided text before issuing the final critique.
Decomposition	If the research problem is complex, **break it down into 4–6 manageable sub-problems** and evaluate the soundness of the solution for each sub-problem individually.

Require the AI to articulate its reasoning process before delivering the final critique. This approach, often using **Chain-of-Thought (CoT)** prompting, significantly enhances performance in reasoning tasks. Specialized techniques like **Chain-of-Verification (COVE)** or **Step-Back Prompting** can be mandated for deeper validation.

3. CORE CRITERIA FOR VALIDITY AND SOUNDNESS (THE DIRECTIVE)

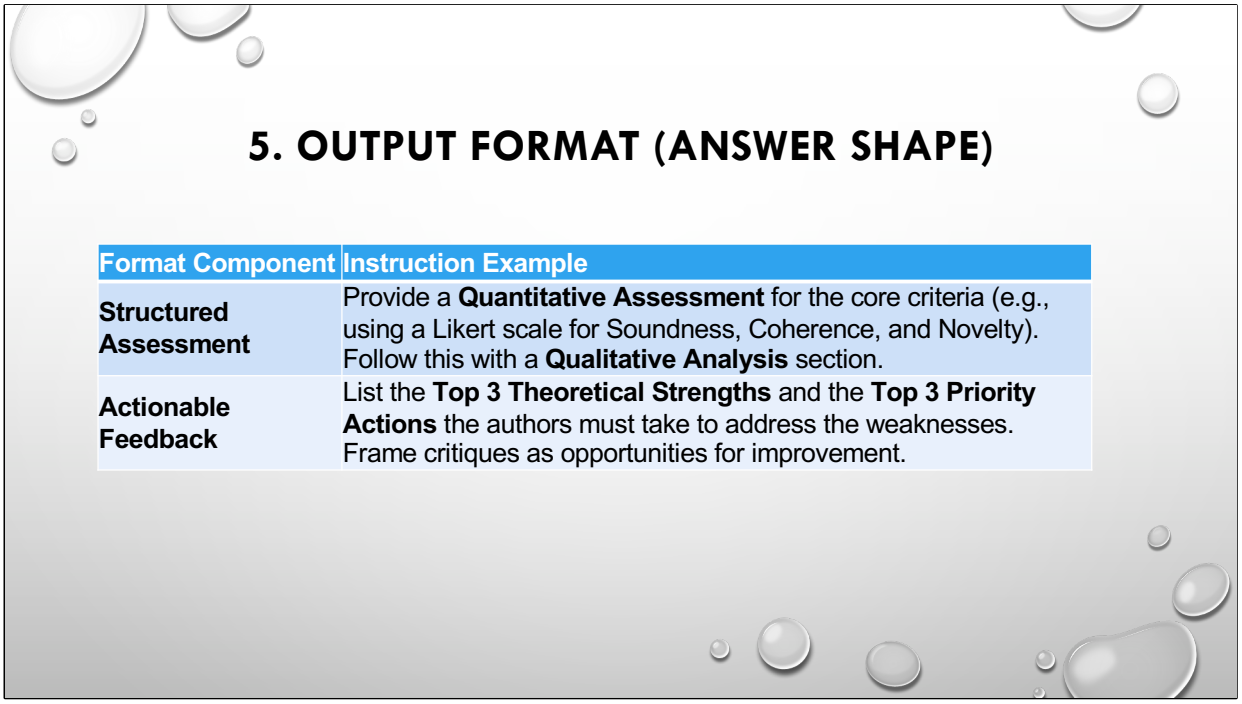
Evaluation Aspect	Instruction Example
Methodological Soundness	Evaluate the methodology for **rigour and reproducibility**. Are the methods **appropriate and well-justified** for the research question? Identify any **omitted quality assurance steps** (e.g., missing method blanks, lack of replicates).
Internal Consistency/Coherence	Assess the logical flow (Structural Organization and Flow). Are all theoretical components **logically aligned without contradictions**? Does the paper contain any **inconsistencies or gaps in the reasoning**?
Evidence & Justification	Evaluate the **statistical validity of the results** and check if the **contribution is novel and justified by the presented evidence**. Spot any instances of **overclaiming** or **cherry-picking results**.
Novelty and Gap Filling	Does the work address a **meaningful problem or gap**? Critically assess whether the work **meaningfully extends beyond existing literature** or addresses gaps/limitations in current understanding.

The core directive must specify the elements of academic soundness the AI must critique, based on established research evaluation criteria.

4. INTEGRATING SELF-CRITICISM AND CORRECTION

Technique	Instruction Example
Active Critique/Self-Refine	**Challenge the paper's central argument.** Identify **potential objections** to its claim and suggest **counterarguments**. Use the provided text to **simulate a peer review critique** to identify weaknesses.
Fact-Checking Mechanism	**Identify all cited references/claims that require external validation.** Due to the high risk of **hallucinations** and **inaccurate references** (up to 70% inaccuracy observed), flag any **fabricated sources or misquotations**.

To overcome the tendency of LLMs to agree too readily or provide generic feedback, instruct the AI to actively look for flaws and generate alternative perspectives.



Specify the format of the output to ensure the critique is actionable and easy for the human researcher to verify.

This comprehensive construct aligns with the concept of **Prompt Engineering**, utilizing roles, detailed instructions, and iterative self-checking mechanisms to push the limits of AI's critical evaluation capabilities, even if the final judgment of validity must be retained by the human researcher.

The integration of Artificial Intelligence (AI) and Generative AI (GenAI) into scholarly communication has led scientific publishers and influential organizations to rapidly develop and update specific policies focusing primarily on authorship, transparency, and integrity.
There is a **broad consensus** among publishers and scholarly organizations regarding the foundational rules for AI use, although specific requirements for disclosure and use during peer review can vary.

KEY PUBLISHER STANCES (SUMMARY)

Publisher	Key Policy Stance
Elsevier	**Prohibits AI authorship**; allows AI only to improve readability/language; **prohibits AI-generated images/artwork**; **prohibits AI use by peer reviewers**.
Springer Nature	**AI does not satisfy authorship criteria**; requires disclosure (e.g., in methodology or acknowledgment); advises reviewers against uploading manuscripts to GenAI tools.
Taylor & Francis	**Bars AI authorship**; mandates disclosure of responsible AI use and human accountability for content integrity.
Science	**Prohibits AI authorship**; requires disclosure of AI use, including **specific prompts** used; **prohibits AI-generated images/multimedia**. Initially banned GenAI use outright but later transitioned to allowing use for writing/presentation with full disclosure.
Wiley	**Allows limited, responsible use** of GenAI, requiring transparency and human accountability; prohibits AI authorship or original data manipulation.
Cambridge University Press	Requires **transparent AI use**; authors are accountable for content integrity and originality; AI lacks authorship.
World Association of Medical Editors (WAME)	AI chatbots should **not be recognized as coauthors**; stresses the necessity for **meticulous disclosure**, including the AI tool's name, version, and specific prompts.

I. AI Authorship and Accountability (Universal Prohibition)

Publishers universally prohibit listing AI tools as authors or co-authors in scientific manuscripts.

Rationale for Prohibition: AI systems cannot be granted authorship status because they **lack the capacity to take moral and legal responsibility** for the submitted work. Authorship implies agency and accountability, which current AI tools do not possess.

Organizational Stance: This prohibition is supported by major organizations such as the **Committee on Publication Ethics (COPE)** and the **International Committee of Medical Journal Editors (ICMJE)**.

Human Responsibility: Human authors must **remain fully responsible** and accountable for the entire content, including the accuracy, validity, and integrity of the results, regardless of AI assistance.

II. Transparency and Disclosure Requirements

Publishers generally agree that the use of AI tools must be transparently disclosed. Disclosure is crucial to uphold the core ethical principles of honesty, transparency, accountability, and the integrity of research.

Mandatory Disclosure (Substantial Use)

Disclosure is **mandatory** when AI use is considered **intentional and**

substantial. AI use is deemed substantial if it **produces evidence, analysis, or discussion** that supports the conclusions, or **directly affects the content** of the research or publication.

Examples of substantial use requiring disclosure include:

Using AI for **data collection, analysis, interpretation, or visualization** (quantitative or qualitative).

Using AI to **formulate questions or hypotheses**, **design and conduct experiments**, or **identify knowledge gaps**.

Using AI to **draft significant parts of the paper**, **summarize, paraphrase, significantly revise, or synthesize textual content**, or to **translate** parts or the whole paper.

Using AI to **generate synthetic data and images** reported in the paper.

Non-Mandatory Disclosure (Trivial Use)

Disclosure is generally **unnecessary** or **optional** for trivial or non-substantial uses that support authors in refining, correcting, formatting, and editing texts and documents, as these uses do not substantially impact accountability or reproducibility.

Examples typically exempt from mandatory disclosure:

Using AI to check basic **grammar, spelling, or references**.

Using AI to suggest words or phrases to enhance **clarity or readability** of an existing sentence.

AI use that is **incidental** to a larger operation where it is **not generating content** (e.g., AI incorporated into internet search engines for background research or laboratory testing machines).

Using AI for **brainstorming** or generating examples for existing content is often deemed **optional**.

Declaration Format

Publishers require detailed disclosure to maintain transparency. Requirements for the declaration typically include:

AI Tool Specification: Identifying the **exact AI tool and version** used.

Specific Role: Describing **how the AI was used** and its specific role in the research process.

Prompts (Science): Science specifically emphasized that the disclosure should include reporting the **specific prompts** used to elicit responses from the AI.

Location of Disclosure:

AI use for writing assistance is often recommended for disclosure in the **acknowledgment section** (e.g., ICMJE).

AI use for methodology, data collection, analysis, or figure generation must be described in the **methods section**.

Some journals require a **new dedicated section** for AI disclosures placed before the references list.

III. Restrictions on Content and Data

Publishers place strong restrictions on the use of AI tools concerning core research elements like data, figures, and results:

Data Integrity: AI is prohibited from being used to **falsify, alter, or manipulate original research data**.

AI-Generated Images/Artwork: Several major publishers have placed **restrictions or outright bans** on the use of AI-generated visual content in publications.

- **Elsevier** prohibits AI-generated images and artwork.
- **Springer Nature** states that journals should not publish AI-generated images, unless analyzing such media or under proper licensing, due to copyright risks.
- The **Science family of journals** bans AI-generated images in manuscripts.

IV. Policies Regarding Peer Reviewers

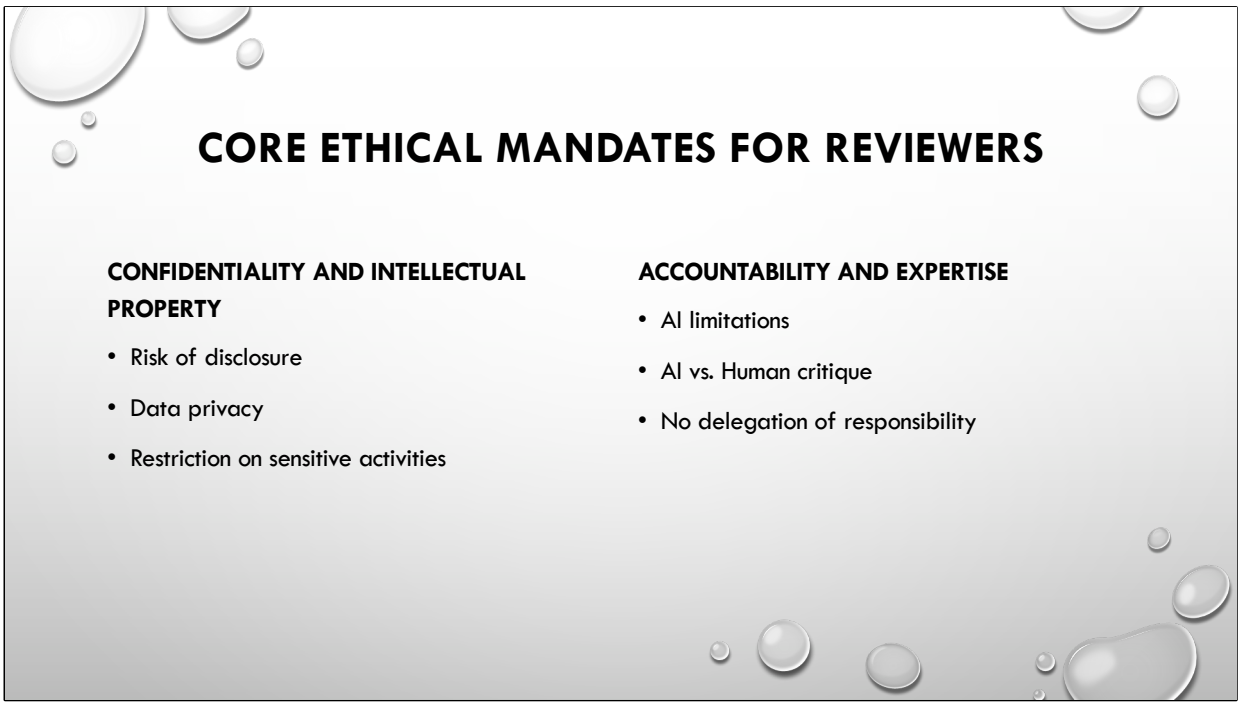
Policies often mandate strict confidentiality and restrict or prohibit AI use by reviewers due to intellectual property concerns and the risks associated with uploading unpublished manuscripts to external GenAI servers.

Elsevier Policy: Elsevier prohibits peer reviewers from using generative AI and AI-assisted technologies in the journal peer review process, even if the use is only intended to improve language and readability. The concern is that uploading manuscript content may violate authors' **confidentiality and proprietary rights** and breach data privacy rights.

Springer Nature Policy: Springer Nature advises peer reviewers not to upload manuscripts into generative AI tools but asks reviewers to **declare transparently** the use of such tools if they supported the evaluation of the claims made in the manuscript.

Funding Agencies: The U.S. National Science Foundation (NSF) and the U.S. National Institutes of Health (NIH) **prohibit** the use of GenAI technologies in the peer review process for grant applications to maintain security and confidentiality.

Rationale: Publishers emphasize that only **human reviewers can be responsible and accountable** for the content of the review report, and critical thinking and original assessment needed for peer review are outside the scope of current AI technology.



Publishers' policies regarding the use of Artificial Intelligence (AI) and Generative AI (GenAI) during the peer review process—often referred to as **Domain 5: Editing, Review, and Publishing Support**—are guided primarily by concerns over confidentiality, accountability, and the integrity of unpublished intellectual property.

Policies vary between strict prohibitions and mandates for transparent disclosure, but all emphasize that human reviewers remain ultimately responsible for the assessment.

Core Ethical Mandates for Reviewers

1. Confidentiality and Intellectual Property

The primary concern is that peer reviewers might upload unpublished manuscripts into external generative AI tools, which violates the authors' intellectual property (IP) rights and compromises the confidentiality of the work.

Risk of Disclosure: Uploading content into generative AI tools that are not secured behind a firewall—like many publicly available models—is considered by funding bodies, such as the **U.S. National Science Foundation (NSF)**, to put that information **in the public domain**.

Data Privacy: This practice may also breach **data privacy rights**, particularly if the manuscript contains personal or proprietary information.

Restriction on Sensitive Activities: General guidelines, such as those from the European Commission's ERA Forum, recommend that researchers **refrain from using generative AI tools substantially in sensitive activities that could impact other researchers or organisations (for example, peer review)**.

2. Accountability and Expertise

Publishers assert that only **human reviewers can be responsible and accountable** for the content of the review report. The critical thinking and original assessment needed for peer review are currently considered outside the scope of AI technology.

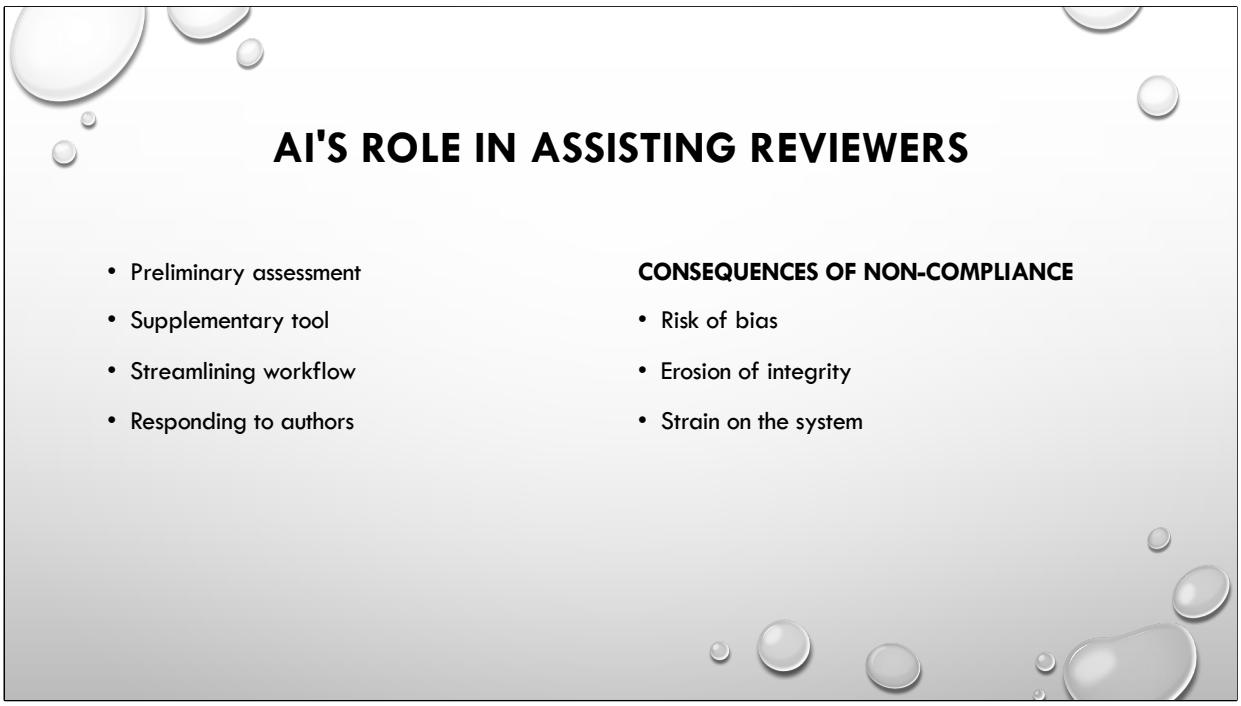
AI Limitations: Reviewers are cautioned because GenAI tools can lack up-to-date knowledge, produce **nonsensical, biased, or false information** ("hallucinations"), or oversimplify complex arguments.

AI vs. Human Critique: Editors can often tell if review comments are AI-generated if they are **relevant but broad in nature** and lack sufficient detail addressing specific issues present in the manuscript, which human experts typically focus on.

No Delegation of Responsibility: Reviewers are ultimately **accountable for the contents** they submit, even if they use AI to assist with their review, and they must fact-check and apply their own judgment to any AI input.

SPECIFIC PUBLISHER AND ORGANIZATIONAL POLICIES

Publisher/Organization	Policy Stance	Details and Rationale
Elsevier	**Strict Prohibition**.	Elsevier **prohibits the use of generative AI and AI-assisted technologies by reviewers** in the journal peer review process. This prohibition extends even to using AI for improving language and readability, due to the risk of violating authors' confidentiality and proprietary rights by uploading manuscript content.
Springer Nature	**Cautious Disclosure/Discouragement**.	Advises peer reviewers **not to upload manuscripts into generative AI tools**. If an AI tool *did* support **any part of the evaluation of the claims made** in the manuscript, the reviewer is asked to **declare the use of such tools transparently** in the peer review report.
Taylor & Francis / Cambridge University Press	**Confidentiality-driven Restriction**.	Requires reviewers to treat manuscripts as confidential and **prohibits uploading files, images, or information from unpublished manuscripts into databases or tools that do not guarantee confidentiality**, including generative AI tools like ChatGPT.
WAME (World Association of Medical Editors)	**Guidance for Attribution**.	Provides recommendations for authors and reviewers on **appropriately attributing chatbot use** in their work.
Science Journals	**Prohibition on AI Images/Multimedia**.	While primarily focused on authors, the journal family has generally stricter stances, including prohibitions on AI-generated images.



Despite prohibitions on uploading confidential manuscripts, AI is seen as having potential in improving the efficiency of the review process, provided the tools are kept safe or the documents are already published:

Preliminary Assessment: AI tools like ChatGPT can provide an "independent opinion" on a manuscript before submission, offering constructive comments on weaknesses and strengths.

Supplementary Tool: Large language models can be used as a supplementary tool to **gain additional perspectives** that might be overlooked in a reviewer's own evaluation.

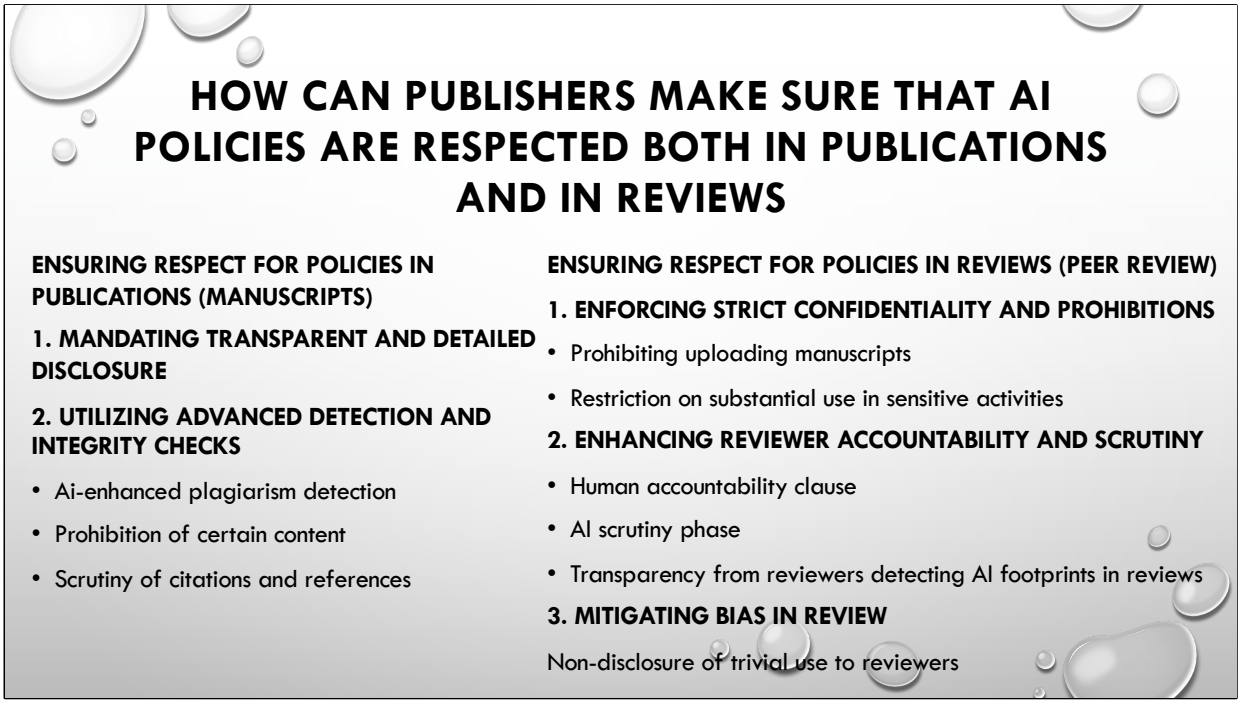
Streamlining Workflow: AI can function as a personal assistant for editors and reviewers, potentially **optimizing the review process** by automating certain repetitive tasks. For instance, AI-based peer review systems can enhance efficiency and objectivity.

Responding to Authors: AI can assist authors in preparing detailed responses to critical comments from peer reviewers, generating point-by-point replies and improving clarity. However, authors must still be vigilant, as AI may cite non-existent references in its responses.

Consequences of Non-Compliance

Publishers and academic bodies underscore the critical risks of non-

compliance, particularly regarding the use of AI during the review phase:
Risk of Bias: Revealing the optional use of AI to peer reviewers could contribute to **biases** and compromise the integrity of the peer review process, possibly leading to biased assessments against non-native English authors.
Erosion of Integrity: If AI-authored material is not detected and rectified during the evaluation process, it **erodes the fundamental integrity** of scholarly work.
Strain on the System: The potential for AI-generated responses from authors puts pressure on editors and reviewers, who must develop strong judgment to ascertain the true source and rigor of the arguments presented.
In summary, publishers are attempting to strike a critical balance by **mandating confidentiality** for peer reviewers—often outright prohibiting the use of external AI tools on submitted manuscripts—while acknowledging that AI *could* be a supportive tool for improving language or providing preliminary critiques, provided the reviewer assumes full, verified accountability for the final output.



The integration of Artificial Intelligence (AI) and Generative AI (GenAI) into scholarly communication has led scientific publishers and influential organizations to rapidly develop and update specific policies focusing primarily on authorship, transparency, and integrity.
There is a **broad consensus** among publishers and scholarly organizations regarding the foundational rules for AI use, although specific requirements for disclosure and use during peer review can vary.

Publishers can ensure that AI policies are respected in both publications (manuscript content) and peer reviews by implementing a multi-layered strategy that combines **mandatory transparency/disclosure requirements**, the deployment of **AI-driven detection technologies**, and the establishment of **clear human accountability standards**.
The goal is to maintain academic integrity, ensure accountability, and preserve confidentiality, particularly in the peer review process, which is considered a sensitive activity.
Here is a detailed breakdown of the mechanisms publishers can use:
I. Ensuring Respect for Policies in Publications (Manuscripts)
Publishers face the challenge of verifying that authors uphold the core principles

of **accountability, transparency, and originality** when submitting AI-assisted manuscripts.

1. Mandating Transparent and Detailed Disclosure

The most critical tool is the process, which is considered a sensitive activity. Here is a detailed breakdown of the mechanisms publishers can use:

I. Ensuring Respect for Policies in Publications (Manuscripts)

Publishers face the challenge of verifying that authors uphold the core principles of **accountability, transparency, and originality** when submitting AI-assisted manuscripts.

1. Mandating Transparent and Detailed Disclosure

The most critical tool is the **requirement for explicit disclosure** of substantial AI use within the manuscript itself.

Mandatory Disclosure Policy: Publishers must define when disclosure is **mandatory**—specifically when AI use is **intentional and substantial** (e.g., when it generates content, analyzes data, or influences conclusions).

Specific Details Required: To prevent vague disclosures, policies must require authors to provide specific details about the AI use:

- **AI Tool and Version:** Specify the exact version of the AI tool used (e.g., GPT-4o).
- **Contribution:** Detail **which parts** were generated or influenced by AI (e.g., data analysis, abstract drafting, or identification of knowledge gaps).
- **Prompts Used: Science** journals, for example, have emphasized that disclosure should include reporting the **specific prompts** used to elicit responses from the AI.
- **Location of Disclosure:** Require disclosure in the **methods section** if AI was used for data collection or analysis, or in a **dedicated section** before the references list for general AI use.

Human Oversight Confirmation: Publishers require authors to confirm that they **reviewed and edited the content** and take **full responsibility** for the entire manuscript to ensure accuracy, coherence, and adherence to ethical guidelines.

2. Utilizing Advanced Detection and Integrity Checks

Since LLMs can produce content that is difficult to distinguish from human writing, publishers rely on technology and expert oversight.

AI-Enhanced Plagiarism Detection: Publishers must invest in **next-level, AI-driven plagiarism detection technologies** that are designed to discern the subtle characteristics and structures unique to AI-generated text, as conventional plagiarism checkers often fail to detect sophisticated AI content.

- **Identifying "Footprints":** Detection systems can look for subtle linguistic "footprints" or phrases common to GenAI models that appear in peer-reviewed scholarly publications, although transparency about use should not be considered an automatic violation.

Prohibition of Certain Content: Publishers can enforce policies that **prohibit** AI-

generated images, figures, or artwork in manuscripts to mitigate concerns about originality, copyright infringement, and data falsification.

Scrutiny of Citations and References: Given that AI frequently generates **inaccurate or hallucinated references** (up to 70% in some cases), publishers must subject all citations—especially those associated with AI-assisted text—to rigorous checking. Editors and reviewers should be vigilant against papers containing falsehoods, fabrications, or biased content resulting from improper AI use.

II. Ensuring Respect for Policies in Reviews (Peer Review Process)

The integrity of the peer review process (Domain 5) is protected by focusing on **confidentiality** and **human accountability**.

1. Enforcing Strict Confidentiality and Prohibitions

Publishers must issue clear rules regarding the use of AI tools on unpublished manuscripts to protect authors' intellectual property and confidentiality.

Prohibiting Uploading Manuscripts: Policies must strongly advise, or outright **prohibit**, peer reviewers from uploading the contents of confidential manuscripts or review reports into external generative AI tools (like ChatGPT). This is critical because uploading content may violate **confidentiality and proprietary rights** and breach data privacy.

- **Elsevier** specifically **prohibits** the use of generative AI by reviewers, even for improving language and readability.
- **Funding Agencies** (like the NIH and NSF) also prohibit the use of GenAI technologies in the peer review process for grant applications to maintain security and confidentiality.

Restriction on Substantial Use in Sensitive Activities: Guidelines recommend that researchers **refrain from using generative AI tools substantially in sensitive activities that could impact other researchers or organizations, such as peer review or evaluation of research proposals**. Substantial use here means delegating the assessment of a paper.

2. Enhancing Reviewer Accountability and Scrutiny

Publishers reinforce accountability by emphasizing that AI cannot replace human judgment and by using systems to detect misuse.

Human Accountability Clause: Publishers must reiterate that **only human reviewers can be responsible and accountable** for the content and conclusions of the review report, as AI tools lack the required critical thinking and moral agency.

AI Scrutiny Phase: Publishers can **integrate an "AI scrutiny" phase** within the peer-review workflow. This involves:

- **Training Reviewers:** Equipping reviewers with training and knowledge about AI applications and their limitations to help them recognize AI-generated content.
- **Preliminary Screening:** Deploying AI algorithms to preliminarily screen

submissions for AI-generated material before they reach human reviewers, which can improve efficiency and objectivity.

Transparency from Reviewers: Policies, such as those from **Springer Nature**, ask reviewers to **declare transparently** the use of any AI tools that supported the evaluation of claims made in the manuscript.

Detecting AI Footprints in Reviews: Journal editors can often identify AI-generated review comments if they are **relevant but broad in nature**, lack sufficient detail regarding specific issues, or exhibit an "unrealistically smooth" and "polished writing" style. Editors must rely on their judgment to ascertain the true source and rigor of the arguments presented.

3. Mitigating Bias in Review

Publishers should consider policies to mitigate bias introduced by AI disclosure in peer review.

Non-Disclosure of Trivial Use to Reviewers: In instances where AI use is **optional or unnecessary** (e.g., basic editing for clarity), publishers might choose to **keep this information from reviewers** to minimize the likelihood of biased assessments, particularly against non-native English authors. This requires further study but addresses the concern that disclosure, even if honest, could inadvertently compromise the integrity of the peer review process.

The integration of Artificial Intelligence (AI) and Generative AI (GenAI) into scholarly communication has led scientific publishers and influential organizations to rapidly develop and update specific policies focusing primarily on authorship, transparency, and integrity.
There is a **broad consensus** among publishers and scholarly organizations regarding the foundational rules for AI use, although specific requirements for disclosure and use during peer review can vary.

The future of research publishing is expected to be **profoundly transformed** by the development of Generative AI (GenAI), leading to a paradigm shift in how scholarly work is created, evaluated, and disseminated. AI is not merely a tool but a **catalyst for transformation** that is set to reshape core scholarly practices and the entire publishing ecosystem.

This transformation is characterized by major changes in efficiency, the nature of intellectual labour, and the necessity for radical policy adaptation.

1. Accelerated, AI-Mediated Workflows

AI is expected to revolutionize and streamline the scientific publishing process across multiple domains, enhancing efficiency and productivity.

Content Generation and Structuring: AI will streamline the writing and editing process, which is integral to **Editing, Review, and Publishing Support**. AI can generate coherent and structured academic prose, assisting in drafting abstracts and summaries, and helping researchers articulate complex data. This capability is expected to accelerate research productivity and benefit scientific advances.

Publication Optimization: AI tools will optimize manuscript preparation by helping authors **align manuscripts with journal guidelines**, **analyze recent publications** to match editorial expectations, and assist in drafting submission materials like abstracts and cover letters.

Data and Analysis Integration: Future multimodal GenAI platforms could enable researchers to input complex datasets, including text, images, video, and audio, and subsequently **generate integrated analyses far more rapidly**, fundamentally changing research processes.
Enhanced Dissemination and Outreach: AI tools are instrumental in **Communication, Outreach, and Ethical Compliance**. AI can transform long-form research into social media-friendly posts, create engaging texts and appealing graphics, and generate summaries for the public, ensuring the work reaches wider audiences with clarity and influence. This streamlines the communication process and supports ethical compliance in research dissemination.

2. A Shift in Authorship and Intellectual Labour

The integration of AI is transforming the nature of scholarly communication and challenging long-standing definitions of academic labour and authorship.
Hybrid Human-AI Writing: The future is moving toward an **"age of hybrid human-AI writing,"** where AI is a collaborative partner in the writing process, reshaping how writing is approached and executed. This evolution suggests that the future scholarly distinction may increasingly depend on **conceptual innovation, theoretical insight, and creative synthesis** rather than traditional writing prowess, as the mechanical aspects of writing are automated.
Collaborative AI Agents: The trajectory of AI development suggests an emerging future where researchers envision AI tools becoming **full collaborative partners**. Emerging **agentic AI models** may be developed to delegate more substantial research tasks to intelligent, goal-driven systems, potentially leading to a "postplagiarism epoch".
Redefined Authorship: This paradigm shift suggests that the evolution of these tools may ultimately **redefine what it means to be an author** in academic contexts.

3. Persistent Threats to Integrity and Accountability

The future of publishing faces serious threats regarding quality, integrity, and trust due to AI's inherent limitations and misuse.
Risk of Fraud and Falsification: AI's ability to generate text and data creates concerns, including the potential for generating **fraudulent scientific articles**. The failure to detect and rectify AI-authored material during the evaluation process **erodes the fundamental integrity** of scholarly work.
Hallucinations and Reliability: AI output is **unreliable** and prone to producing **inaccurate or falsified information**. This unreliability necessitates extensive fact-checking by humans, and improper AI use could lead to an **increased rate of retractions** and potentially disrupt subsequent research.
Deskilling and Stagnation: Reliance on AI is linked to the risk of **diminishing critical thinking** and **academic creativity**. AI's tendency to rely on past patterns risks leading to a **stagnation of academic discourse**, where researchers recycle

prevailing perspectives rather than generating new insights.

AI in Peer Review: AI usage may extend to the peer review process, potentially leading to **GenAI-produced manuscripts being reviewed by GenAI tools themselves**. This raises concerns about compromised scholarly rigour and the integrity of the process.

4. The Policy Imperative

The increasing influence of GenAI necessitates a continuous and collaborative policy response from publishers and academic bodies.

Mandatory Transparency: Publishers are increasingly requiring **transparency and detailed disclosure** of substantial AI use in manuscripts. Policies are constantly evolving to provide clarity on **how** AI should be used and disclosed. For example, the journal *Science* transitioned from an outright ban on GenAI use to allowing its use for writing or presentation, provided there is **full disclosure**.

Enhanced Detection Tools: There is an **urgent demand** for publishers to develop and implement **more robust, AI-focused plagiarism detection systems** capable of adapting to the advanced capabilities of evolving AI writing techniques.

Focus on Ethical Frameworks: Future research and industry efforts must focus on establishing ethical frameworks for AI use, defining the appropriate boundaries for human-AI collaboration models, and providing adequate **training for researchers** to ensure ethical and transparent application.

The future of research publishing will therefore be defined by the necessity to **balance the enormous productivity gains offered by AI** with the **ethical imperative to maintain human oversight, transparency, and scholarly integrity**.

The advancement of Generative Artificial Intelligence (GenAI) is viewed as a **pivotal moment** that will continue to have a **profound impact on research**. Experts believe that AI technologies are expected to **profoundly transform scientific research, publication, science communication, and research evaluation**. The integration of AI into academic endeavors is considered an **unavoidable progression**, establishing AI as a **catalyst for transformation** that reshapes scholarly thinking, institutional practices, and ethical debates.

Here is a detailed description of the future of research in light of AI advancement, drawing upon the identified domains and challenges:

1. Accelerated, AI-Mediated Research and Discovery

The future of research is characterized by unparalleled efficiency and the ability to pursue fundamentally new types of inquiry.

Boosting Productivity and Efficiency: AI is expected to accelerate **research productivity** and benefit **scientific advances in multiple fields**. Research tasks that previously took significant time and effort are performed in seconds. This efficiency allows researchers to focus on **groundbreaking ideas and intricate problem-solving**.

Transforming Research Domains: AI is revolutionizing research across all six core domains of academic work. This includes **idea development and research**

design, **literature review and synthesis**, and **data management and analysis**.
Uncovering Novel Insights: AI's advanced data analysis capabilities enable it to **uncover patterns and correlations that might be invisible to human researchers**. AI tools can process and synthesize vast amounts of information from different disciplines, helping researchers to utilise findings from other fields, leading to **innovative solutions**. This capacity can **reveal novel applications of existing knowledge**.
Enabling Interdisciplinary Research: AI holds immense potential to **revolutionise and streamline interdisciplinary research**, acting as a bridge between diverse fields. As a general-purpose technology, AI facilitates **interdisciplinary innovation** by encouraging integrated approaches to global challenges.
Agentic and Multimodal Systems: Future advancements point toward AI becoming **more autonomous, multimodal, and seamlessly integrated** into academic workflows. Emerging **agentic AI models** are being developed to delegate **more substantial research tasks** to intelligent, goal-driven systems. Multimodal GenAI platforms could enable researchers to input complex datasets (including text, images, video, and audio) and subsequently **generate integrated analyses far more rapidly**.
Predictive Analysis: Future research directions include the development of **advanced AI tools for hypothesis formulation and predictive analysis**. AI can analyze large volumes of data to **identify emerging research trends and predict future directions**.

2. The New Paradigm of Human-AI Symbiosis

The future research enterprise will be defined by collaboration, requiring a re-evaluation of human expertise and academic identity.
Collaborative Partnership: The future of research lies not in human versus AI competition, but in **thoughtfully designed partnerships** that amplify the best of both human and artificial intelligence. AI is expected to become a **full collaborative partner**.
Symbiotic Scholarship: This new paradigm, referred to as **"symbiotic scholarship,"** requires that human creativity, ethical reasoning, and contextual understanding **combine synergistically** with AI's computational power and pattern recognition capabilities to address complex challenges.
Shifting Intellectual Focus: Since the mechanical aspects of writing and data wrangling may be increasingly automated, scholarly distinction will increasingly depend on **conceptual innovation, theoretical insight, and creative synthesis** rather than traditional writing prowess.
Human Agency is Central: The role and importance of the researcher are becoming **more central** to research, rather than less so. Human scientists are well-positioned to be **early adopters of AI** due to their critical and analytical thinking. AI is envisioned as a collaborator and amplifier of human potential, with

frameworks developed to **safeguard human creativity and agency**.

3. Structural Risks and Threats to Academic Integrity

Despite the opportunities, the accelerated adoption of AI poses significant ethical and integrity challenges that must be continually addressed.

Erosion of Integrity and Falsification: AI poses significant risks of **falsification of information** and the potential for machines to **generate fraudulent but authentic-looking scientific medical articles**. The failure to detect and rectify AI-authored material during the evaluation process **erodes the fundamental integrity of scholarly work**.

Deskilling and Indolence: Improper reliance on AI can lead to the **deskilling** of researchers by progressively eroding their ability to perform essential research tasks, such as generating ideas and conducting literature reviews. This reliance fosters **academic indolence** (laziness), where scholars become relaxed and carefree because AI software can produce up to 90% of their paper.

Diminution of Critical Thinking: Relying on AI limits **critical thinking**, academic creativity, and originality in scholarship. AI machines **do not think** and therefore cannot reflect, merely delivering output as programmed.

Algorithmic Bias: AI models, which are built on vast datasets, can **perpetuate unfairness and discrimination** if the training data contains biases. This tendency risks **reinforcing existing academic hierarchies** and marginalizing alternative voices.

Loss of Autonomy: Improper AI-assisted research ceases to be an autonomous activity and becomes a **heteronomous one, fundamentally guided and directed by AI algorithms**. By relying on unsupervised AI, researchers risk losing ownership of their own research and relinquishing their expertise and judgment.

The Paradox of Ethical AI-Assisted Research: If AI is used ethically and responsibly, requiring human oversight and verification of AI outputs (such as literature reviews or idea generation), it **does not provide significant net benefits in terms of increased efficiency and time-saving** for primary research tasks.

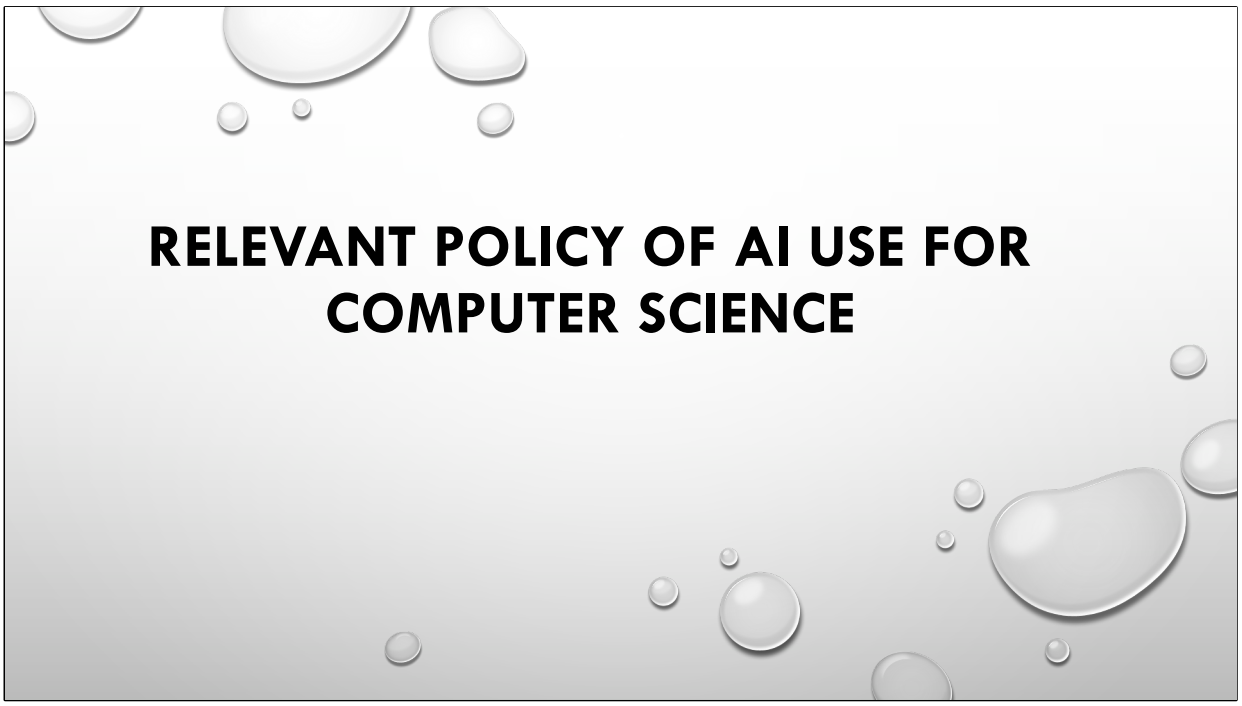
4. The Policy and Educational Imperative

The future requires a concerted effort from institutions and policymakers to navigate these challenges through training, governance, and a commitment to equity.

Need for Ethical Frameworks: The rapid adoption of GenAI has **outpaced the development of ethical guidelines**. There is a pressing need for **adequate training and adaptation periods** for researchers to become proficient in using AI tools effectively. These ethical frameworks must address challenges such as **data privacy, algorithmic bias, and maintaining academic integrity**.

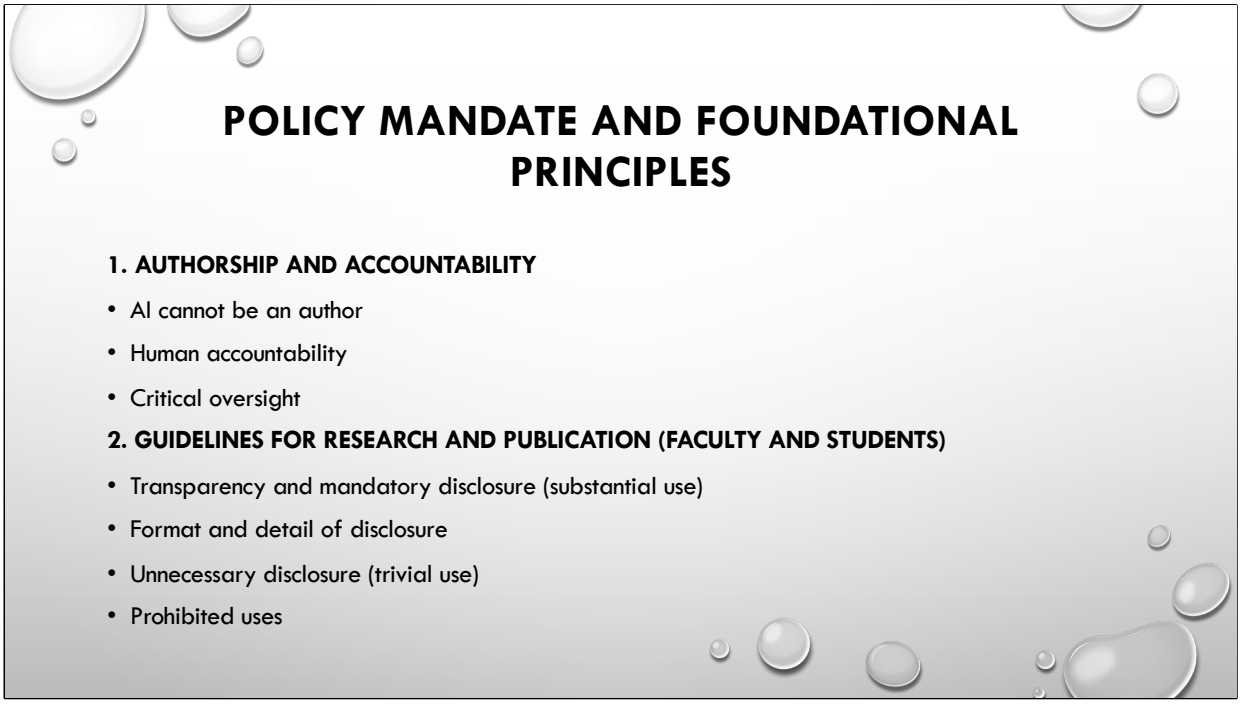
Closing the AI Education Gap: The impact of AI on scientific disciplines has **raced ahead** across science, but the educational focus on AI to upskill future scientists has **lagged**. There is a **misalignment between the impact of AI and AI**

education.
Equity and Inclusion: AI's growing impact on science may **further exacerbate existing inequalities** in science, as disciplines with a higher proportion of women or Black scientists tend to be associated with **less benefit** from AI engagement. This underscores the importance of **expanding participation and opportunities** in AI research for underrepresented groups.
Continuous Monitoring: The landscape of AI in academia is **ever-evolving**, necessitating **continued research and development** to refine tools, address emerging challenges, and consider ethical implications.
The future of research is thus akin to navigating a complex terrain with a powerful, fast, but sometimes unpredictable vehicle: the potential for speed and discovery is immense, but the driver (the human researcher) must maintain **unwavering vigilance, ethical judgment, and ultimate control** to ensure a successful and trustworthy journey.



This policy outlines the principles and procedures for the ethical, transparent, and responsible utilization of Artificial Intelligence (AI) and Generative AI (GenAI) tools within the Faculty of Computer Science (CS) at [Institution Name]. Given that Computer Science is a discipline with one of the largest AI impact scores, and its core involves the application and selection of algorithms, this policy emphasizes balancing innovation with stringent ethical compliance.

Sample policies and guidance at other universities
Boston University – Using Generative AI in Coursework
Stanford – Generative AI Policy Guidance | Office of Community Standards
USC – CIS-Generative-AI-Guidelines



Policy Mandate and Foundational Principles

The purpose of this policy is to maximize the benefits of AI as an essential **productivity tool** while ensuring that the core values of academic integrity—**Authenticity, Transparency, and Accountability**—are upheld in all research, teaching, and administrative functions.

1. Authorship and Accountability

1.1. **AI Cannot Be an Author:** AI tools, including Large Language Models (LLMs) like ChatGPT, Claude, or Gemini, **cannot be listed as authors or co-authors** on any scholarly output, research proposal, or assessment. Authorship implies the capacity to take **moral and legal responsibility** for the submitted work, which AI systems lack.

1.2. **Human Accountability: Human researchers remain ultimately responsible** and accountable for the entire content of the work, including the accuracy, originality, ethical compliance, and integrity of the results, regardless of AI assistance.

1.3. **Critical Oversight:** Users must maintain a **critical approach** to the output produced by GenAI, remaining aware of the tools' limitations, such as **bias, hallucinations, and inaccuracies**. Outputs must be **carefully reviewed and verified** before inclusion.

2. Guidelines for Research and Publication (Faculty and Students)

Research in the Faculty of Computer Science utilizes AI across all six recognized domains of academic function. Transparency is the mandatory principle guiding AI use in research outputs.

2.1. Transparency and Mandatory Disclosure (Substantial Use)

Disclosure of AI use is **mandatory** when the application is **intentional and substantial**. AI use is substantial when it produces evidence, analysis, or discussion that supports the conclusions/findings of a study, or directly affects the content of the research or publication.

Mandatory disclosure is required when AI is used for the following functions in Computer Science research:

Idea Development and Research Design (Domain 1): Using AI for **hypothesis generation**, **identifying literature gaps**, or assisting in **study design**.

Literature Review and Synthesis (Domain 3): Using AI for **automated literature synthesis**, extracting or **analysing vast volumes of information** via text extraction or semantic analysis.

Data Management and Analysis (Domain 4): Using AI for **data analysis descriptions, detailed analysis, visualization**, or **dataset management** (creation and curation). This includes using AI to **analyze genomic data, text, or radiologic images**.

Content Generation (Domain 2 & 5): Using AI to **significantly draft, summarize, paraphrase, or synthesize textual content** (beyond basic editing).

Code Generation: Using AI tools (e.g., specialized coding assistants) to **generate substantial portions of code** used in the research methodology or results.

2.2. Format and Detail of Disclosure

When disclosure is mandatory, the human author must report the usage transparently:

Location: If AI was used for data collection, analysis, or figure generation, the use **must be described in the Methods section**. General writing assistance should be acknowledged in the **acknowledgment section**.

Detail Required: The disclosure must include:

- The **specific AI tool and exact version** used (e.g., OpenAI GPT-4-Turbo, Feb. 2024).
- The **specific contributions** AI made (e.g., "Drafted the introduction section," "Analysed the network data").
- The **data and prompts used** to elicit responses from the AI.

2.3. Unnecessary Disclosure (Trivial Use)

Disclosure is **unnecessary** when the AI use is considered basic or non-substantial. These uses are generally seen as basic author support tools and include:

Using AI to **copyedit a paper for grammar, spelling, or suggest synonyms** or

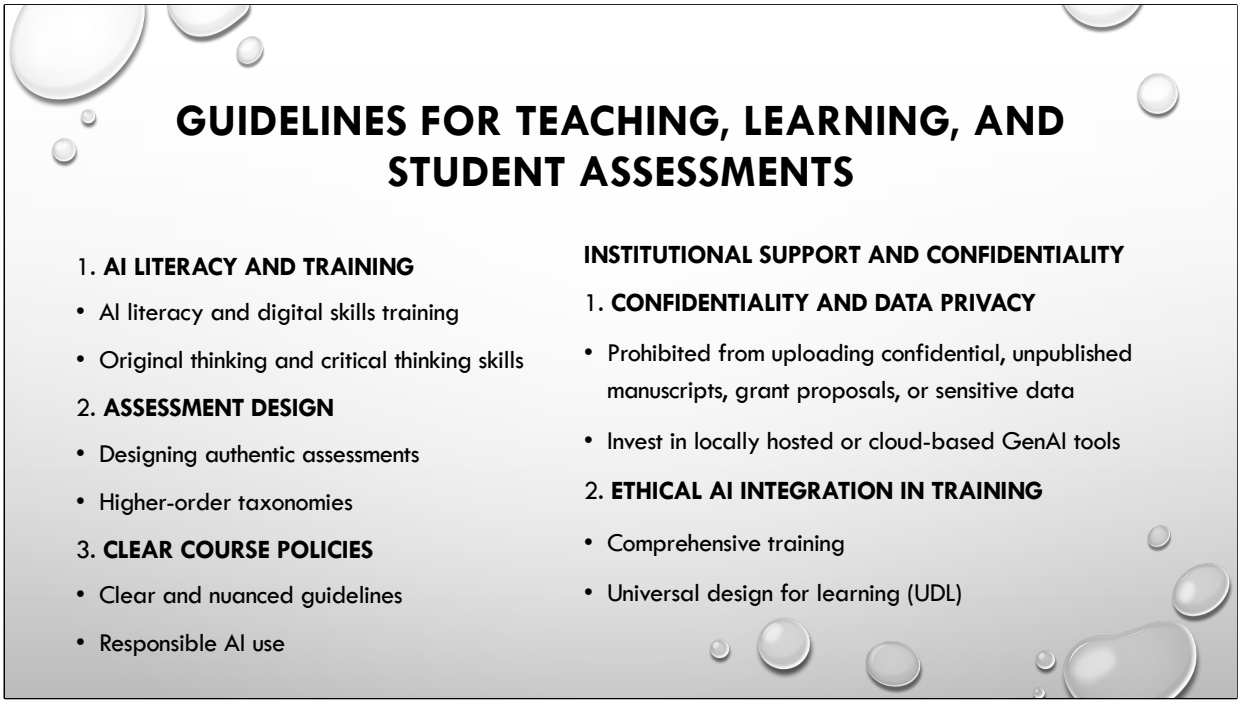
improve existing phrases.
Using AI that is **incidentally integrated into other standard software** (e.g., AI components in internet search engines or word processors).

2.4. Prohibited Uses

Data Integrity: AI must **not be used to falsify, alter, or manipulate original research data**.

AI-Generated Images: Generating images, figures, or multimedia for research publications with AI is **prohibited** unless the AI-generated media itself is the subject of the research.

AI Plagiarism (AIgiarism): Misrepresenting AI-generated content as original work, or **relying on AI to paraphrase without proper citation**, constitutes **academic misconduct**. Authors must **substantially rephrase** AI-generated content in their own words and properly cite sources.



Guidelines for Teaching, Learning, and Student Assessments

The Faculty of Computer Science must focus on fostering a culture of ethical AI use in education and reframing assessment in light of GenAI's capabilities.

3.1. **AI Literacy and Training:**

The Faculty must integrate **AI literacy and digital skills training** into the undergraduate and postgraduate curriculum. This training must cover the ethical dimensions of AI, including **bias mitigation, data privacy, and intellectual property rights**.

Training should focus on teaching students how to use AI as a supplementary tool—not as a substitute for **original thinking** and **critical thinking skills**.

3.2. **Assessment Design:**

Educators are responsible for **designing authentic assessments** that cannot be easily fulfilled by AI alone.

Assessment should focus on **higher-order taxonomies** and prioritize evaluating **sophisticated thinking capabilities, originality, and the process** rather than just the final product.

3.3. **Clear Course Policies:**

Faculty must provide students with **clear and nuanced guidelines** on the permissible and non-permissible use of AI tools for every assessment, and these

guidelines **should not be silent** in the course syllabus.
Instructors should consider modeling **responsible AI use** in the classroom and providing **scenario-based dilemmas** to cultivate ethical decision-making skills in students.

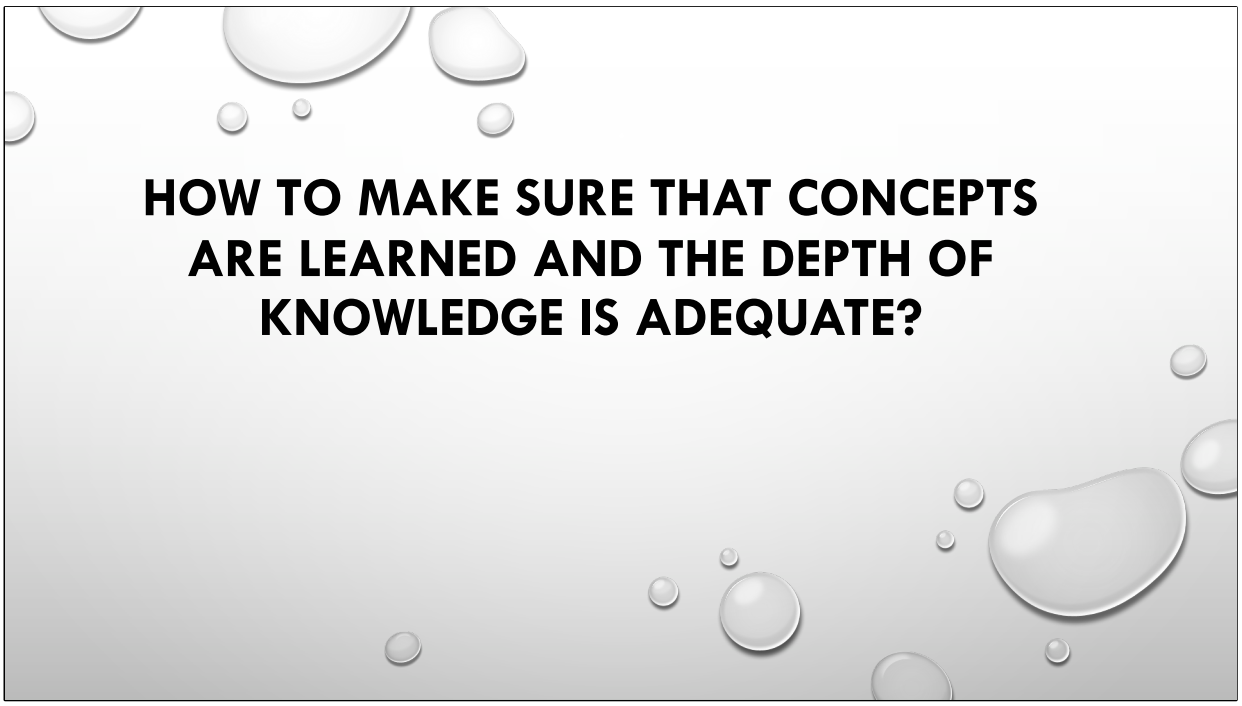
4. Institutional Support and Confidentiality

4.1. **Confidentiality and Data Privacy:**

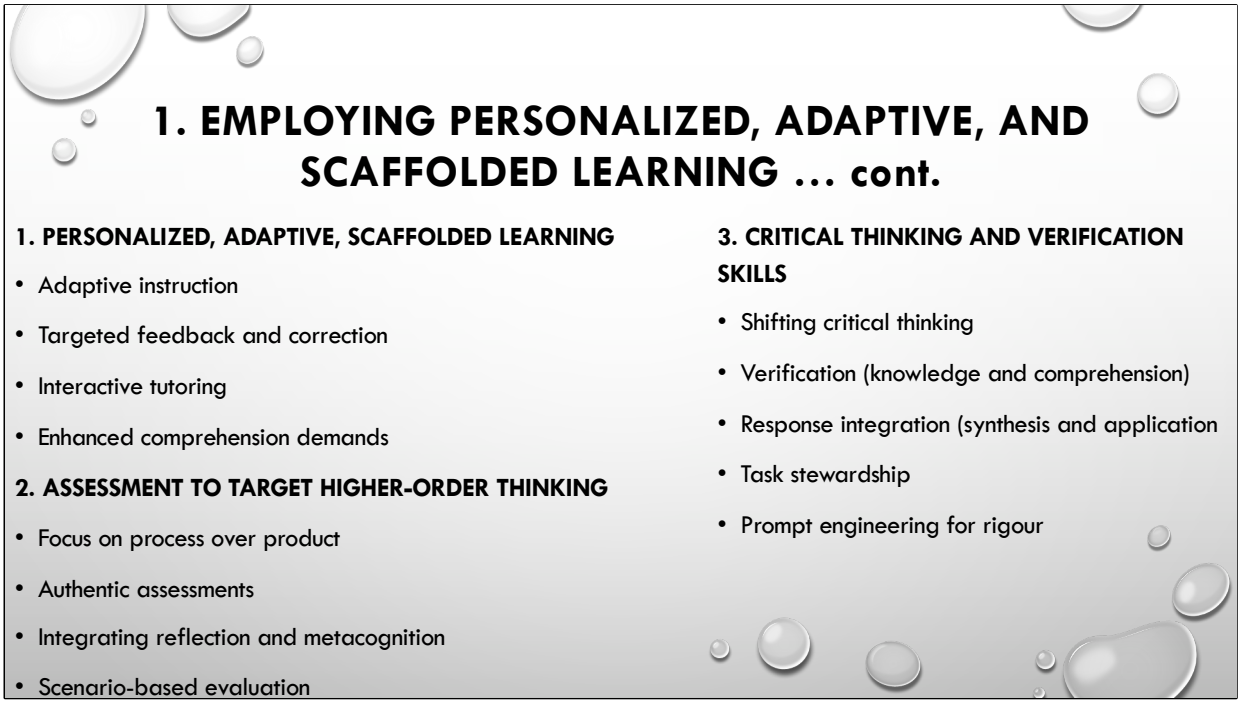
Faculty and students are **prohibited from uploading confidential, unpublished manuscripts, grant proposals, or sensitive data** (including personal student data) to non-approved external generative AI tools that do not guarantee confidentiality. Uploading such information to external tools not behind the institutional firewall is deemed to put that information in the public domain.
The institution should **invest in locally hosted or cloud-based GenAI tools** that it governs itself to ensure data protection and confidentiality for research activities.

4.2. **Ethical AI Integration in Training:**

The institution must provide **comprehensive training** for staff at all career levels on using GenAI, focusing especially on verifying output, **maintaining privacy, addressing biases**, and protecting IP rights.
The Faculty should embrace AI tools that support **Universal Design for Learning (UDL)**, recognizing AI's potential as an accessibility tool to support social justice in higher education.



Ensuring that concepts are learned and the depth of knowledge is adequate in the age of AI requires educational and research strategies that move beyond traditional assessment and passive information consumption. The shift involves designing learning experiences and assessments that cultivate **critical thinking, metacognition, transparency, and human-AI collaboration**, focusing heavily on the **process** of knowledge acquisition rather than just the final output.
Here are detailed strategies, drawing on pedagogical theories and emerging AI practices:



1. Employing Personalized, Adaptive, and Scaffolded Learning
AI tools can serve as a **more knowledgeable other (MKO)** to provide tailored support that ensures students operate within their **Zone of Proximal Development (ZPD)**, leading to deeper understanding and better retention.
Adaptive Instruction: AI platforms dynamically assess a learner's abilities and pace, adjusting content difficulty and providing scaffolding that is **challenging yet achievable**. This ensures the support is precisely within the student's cognitive zone, promoting engagement and motivation.
Targeted Feedback and Correction: AI tools offer **real-time feedback** on assignments, quizzes, and practice exercises, helping students understand and learn from their mistakes immediately. This support progresses the learner from dependence to **autonomy**.
Interactive Tutoring: Large Language Models (LLMs) can **mimic one-to-one tutoring** and instructional dialogue, assisting students' comprehension and retention. Students can engage in **negotiation of meaning** with the AI, querying and questioning responses to refine their understanding of complex concepts.
Enhanced Comprehension Demands: For challenging materials, especially in academic writing or non-native language contexts, GenAI can assist with **summarizing and categorizing literature** or translating content, addressing

comprehension challenges and allowing students to focus on deeper analysis.

2. Redesigning Assessment to Target Higher-Order Thinking

Traditional assessments are easily exploited by GenAI and often fail to measure depth of understanding. Adequate knowledge depth must be evaluated through tasks that require **sophisticated thinking capabilities**, which GenAI cannot easily fulfill.

Focus on Process over Product: Assessments should shift focus to the **process as opposed to the product of learning**. This can involve tracking students' cognitive and problem-solving development over time using AI-assisted formative assessment.

Authentic Assessments: Designing assessments that are **authentic** (real-world problem-solving) and require **human-unique capabilities** like creativity, emotional intelligence, and critical thinking is crucial. Examples include **portfolio-based assessments** demonstrating real-world problem-solving.

Integrating Reflection and Metacognition: Assessments should require students to demonstrate **metacognition** (the ability to understand and regulate one's own thinking) and **self-regulated learning**. This can be achieved by making AI interaction an explicit part of the assignment, requiring students to **reflect on their dialogue with the AI** and justify their decisions.

Scenario-Based Evaluation: Using **scenarios and dilemmas** that involve complex "grey areas" helps cultivate **ethical decision-making skills**, ensuring students understand the boundaries and context of applying their knowledge.

3. Cultivating Critical Thinking and Verification Skills

The core of ensuring adequate knowledge depth lies in teaching students to critically evaluate AI output, acknowledging that AI is based on statistical patterns and is prone to generating errors or "hallucinations".

Shifting Critical Thinking: GenAI shifts the nature of critical thinking toward **information verification, response integration, and task stewardship**. Training must focus on developing these skills.

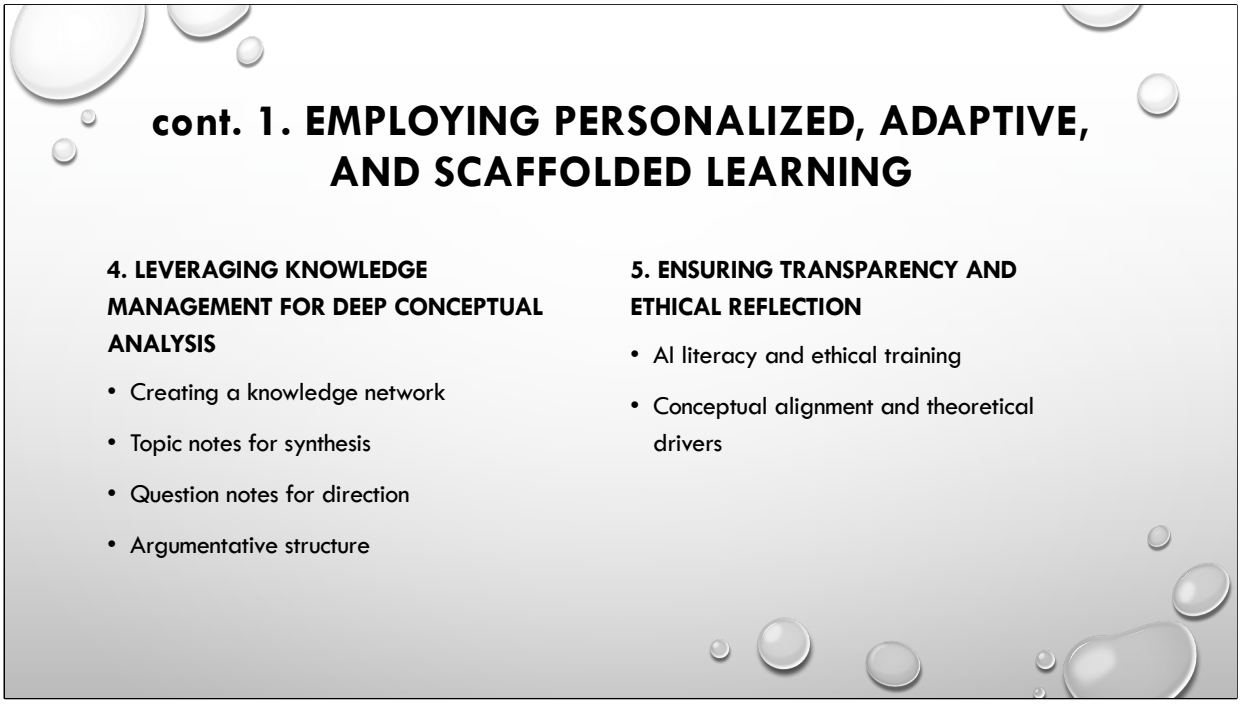
Verification (Knowledge and Comprehension): Students must be trained to **verify AI output against independent standards or through empirical testing**. This addresses the risk that AI output can be wrong and needs verification. AI tools themselves can be prompted to articulate the data that requires external validation.

Response Integration (Synthesis and Application): Critical thinking is enacted when students assess how well AI output meets **objective criteria** or **subjective standards** (e.g., logical flow, feasibility, relevance). Students must learn to **critically customize and refine** AI-generated content and integrate **only relevant parts** into their work.

Task Stewardship: Students must maintain a reflective approach, ensuring they set clear goals, refine prompts, and consciously decide how and when to use AI. This requires students to actively **push back against the AI's affirmations** and

instruct it to critique their ideas or identify counterarguments to strengthen their work.

Prompt Engineering for Rigour: Researchers and students must master prompt engineering, as **better prompts lead to improved results**. This includes using techniques like **Chain-of-Thought (CoT)**, which encourages the LLM to express its reasoning process before delivering its final answer, making it easier for the human to check for logical consistency.



4. Leveraging Knowledge Management for Deep Conceptual Analysis
The development of **academic writing** involves deeply analysing concepts, which leads to a well-explained theory or conclusion, thus ensuring thorough understanding. Academic Knowledge Management systems, especially those focusing on linking concepts, support this depth.
Creating a Knowledge Network: Depth of knowledge is achieved through the **connection of ideas**. Students should consistently **link new notes to existing notes** to ensure their knowledge network is well-connected. This creates a **knowledge graph** that can be traversed to "harvest" new ideas and insights.
Topic Notes for Synthesis: Students should create **Topic Notes** which synthesize information from **multiple sources** around a single concept or theme and are written in the student's **own words and understanding**. This process forces internalization and reflection, reducing ambiguity and confusion.
Question Notes for Direction: Students should cultivate **Question Notes** to capture and refine **research questions, ideas, or insights**. As the original research is built around answering specific research questions, developing these notes ensures focused intellectual direction and identifies potential solutions.
Argumentative Structure: Students can use digital canvases to visually break down arguments, ideas, or projects into a structured sequence of **hypotheses,**

questions, answers, and explanations, allowing for the exploration of concepts and the development of a clear plan.

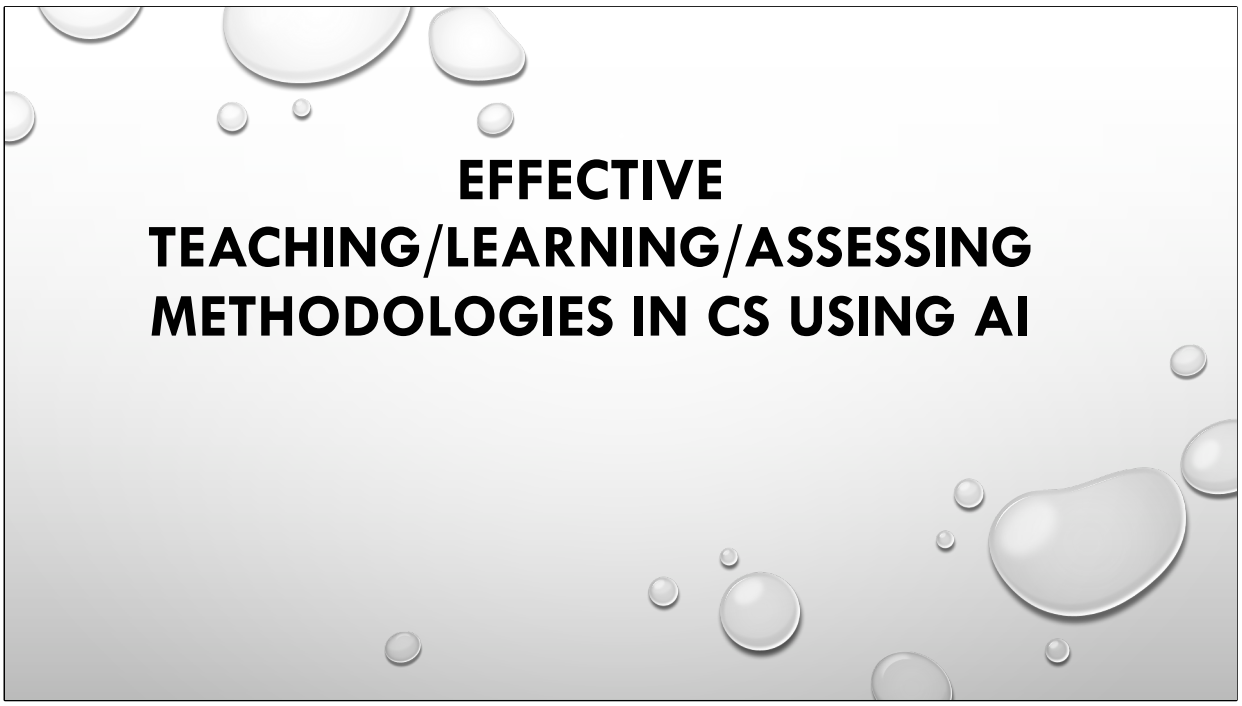
5. Ensuring Transparency and Ethical Reflection

Adequate learning includes developing a profound **philosophical understanding of research** and the ethical implications of using AI, which directly impacts the quality and trustworthiness of the work.

AI Literacy and Ethical Training: Academic institutions must provide **AI literacy training** that teaches students to use GenAI **ethically and responsibly**. This training should address topics like AI bias, data privacy, and the importance of **transparency**.

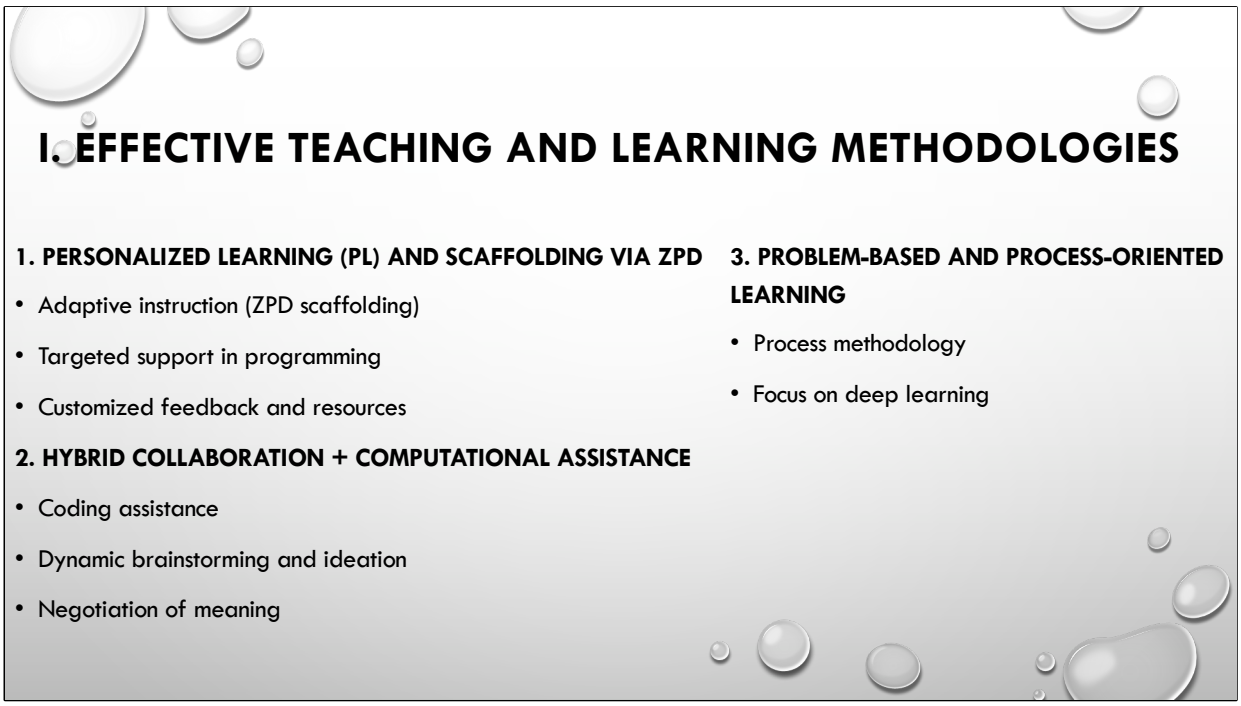
Conceptual Alignment and Theoretical Drivers: Students must be encouraged to understand the **philosophical world view** and **conceptual underpinnings** of their research, ensuring their work is a systematic investigation based on existing theories. This is crucial for strengthening the argument about the relevance and value of their research. Asking specific reflective questions helps students ensure their results challenge or support existing theories.

By implementing these multi-faceted strategies, institutions can shift the focus from rote learning and submission avoidance toward **active knowledge construction, critical application, and responsible collaboration**, ensuring students develop the necessary depth of knowledge for their field.



The rapid integration of Generative AI (GenAI) necessitates that Computer Science (CS) faculties adapt their pedagogical approaches, learning experiences, and assessment frameworks. These new methodologies must focus on leveraging AI's capacity for efficiency and personalization while simultaneously emphasizing human critical thinking, transparency, and ethical accountability.

Effective teaching, learning, and assessment methodologies in Computer Science using AI tools draw heavily on concepts like **Personalized Learning**, the **Zone of Proximal Development (ZPD)**, **Hybrid Collaboration**, and **Reframed Assessment**.



AI tools are essential productivity tools that support six core domains of academic work. In a technical discipline like Computer Science, AI enhances learning by acting as a personal assistant, a perpetual tutor, and a collaborative coding partner.

1. Personalized Learning (PL) and Scaffolding via ZPD

AI's most significant advantage is its ability to tailor educational experiences to individual students' needs, strengths, weaknesses, and interests.

Adaptive Instruction (ZPD Scaffolding): AI tools function as a "more knowledgeable other" (MKO) by providing **scaffolding** that helps students transition from what they can do independently to what they can achieve with guidance. AI tools dynamically assess a student's current level and **adjust content difficulty in real-time**, ensuring that instruction aligns precisely with their cognitive zone.

Targeted Support in Programming: This is highly relevant in CS education for programming and complex technical concepts. AI can provide **real-time feedback, adaptive guidance, and interactive learning experiences**, mimicking a human tutor. Tools can analyze a student's past performance to identify difficulties and provide **targeted support** in those areas. For instance, chatbots have been explored as accessible tools to improve learning

engagement among programming students, even those **without prior coding experience**.

Customized Feedback and Resources: AI generates personalized recommendations for educational resources, practice questions, exercises, and learning paths based on a student's performance, interests, and goals. This customized feedback leads to a more individualized and effective learning experience.

2. Hybrid Collaboration and Computational Assistance

AI tools must be positioned as collaborators that enhance, rather than replace, human critical thinking and creativity.

Coding Assistance: For CS students, Large Language Models (LLMs) **transform coding into an intuitive, conversational process**, making it more accessible and efficient for both novices and experts. AI offers various types of coding assistance:

- **Code Generation:** Generating substantial portions of code.
- **Code Understanding:** Providing a high-level overview by highlighting structure and explaining the logic behind specific implementations, which is crucial for interpreting legacy code.
- **Debugging and Optimization:** Assisting in identifying errors and improving the efficiency of code.
- **Learning and Translation:** Facilitating the acquisition of new coding practices.

Dynamic Brainstorming and Ideation: Students should use AI for **dynamic brainstorming and ideation** to refine writing styles, develop arguments, or explore perspectives outside their immediate expertise (e.g., asking the AI to role-play an HR professional). This supports **idea generation and academic project planning**.

Negotiation of Meaning: Effective learning involves continuous engagement in an iterative cycle of feedback and **negotiation of meaning** with the AI, where students seek understanding and correct misunderstandings. This interactive process helps students refine their understanding of complex concepts and produce better linguistic output, akin to acquiring a new academic "language".

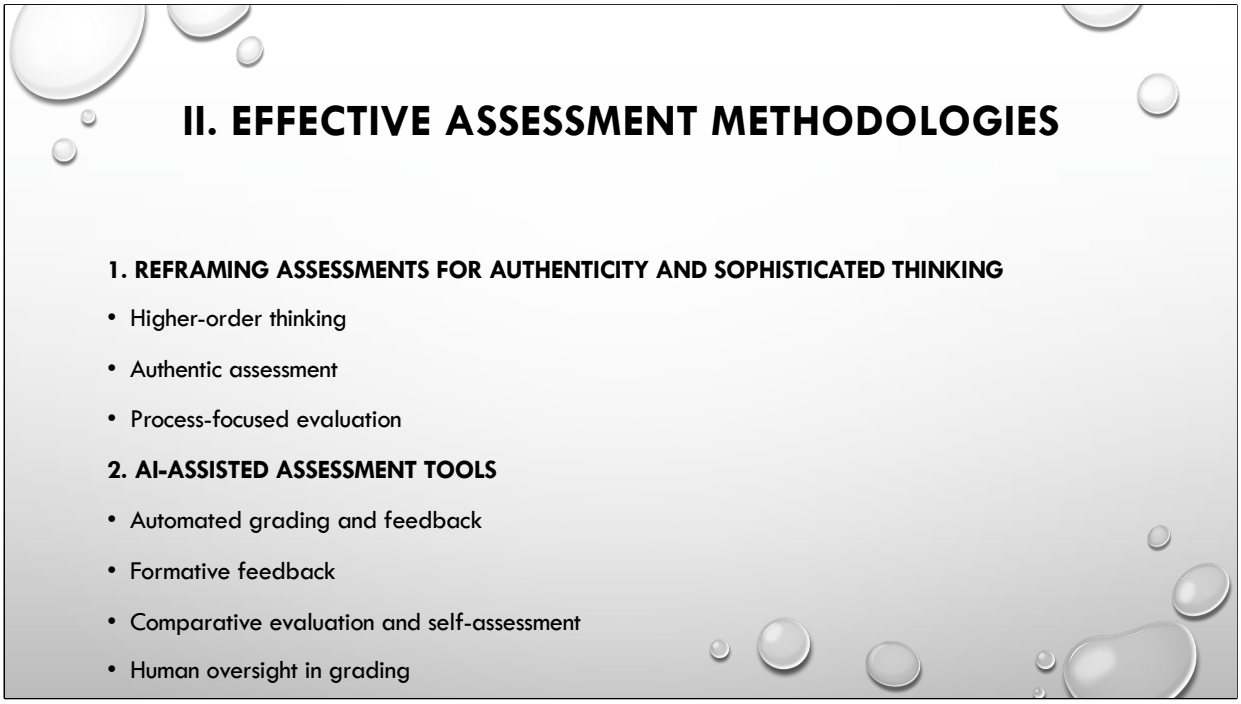
3. Problem-Based and Process-Oriented Learning

Given that CS involves the ability to select appropriate algorithms and methodologies, teaching must prioritize the underlying research processes.

Process Methodology: In Computing Science, process methodologies are used to **understand the processes** involved in tasks like the design and construction of software systems, human-computer interactions, and cognitive modeling in the field of AI.

Focus on Deep Learning: Assessment should move beyond the final product to focus on the **process of learning**. Problem-based learning fosters **deep learning, critical thinking, and collaborative skills**, which are essential when

leveraging AI's analytical power.



Traditional assessment practices, which often fail to capture diverse knowledge, can be exploited by the misuse of GenAI tools. Therefore, assessments must be fundamentally redesigned.

1. Reframing Assessments for Authenticity and Sophisticated Thinking

Higher-Order Thinking: Assessments must prioritize the evaluation of **sophisticated thinking and learning capabilities**, such as **metacognition and self-regulated learning**. The use of AI in education is generally constrained by requiring students to engage with **higher-order taxonomies**.

Authentic Assessment: Educators must design **authentic assessments** that cannot be easily fulfilled with AI. Authentic assessment should integrate GenAI tools while mitigating academic dishonesty and enhancing the learning experience. Examples include portfolio-based assessment demonstrating **real-world problem-solving**.

Process-Focused Evaluation: Assessments need to shift focus to the **process as opposed to the product of learning**. For instance, requiring students to document their iterative dialogue and refinement process with AI, or submitting reflexive journals, can help evaluate the critical journey rather than just the final output.

2. AI-Assisted Assessment Tools

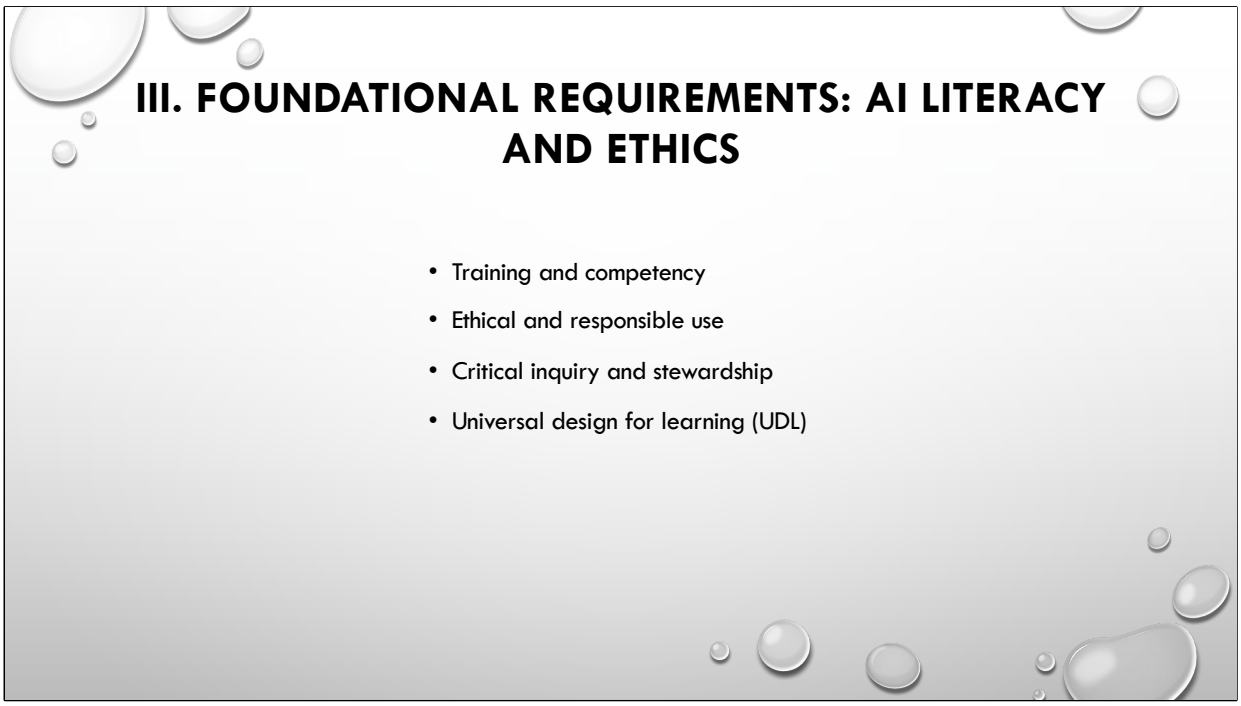
AI tools can assist educators in grading and providing feedback, but they must be managed with human oversight.

Automated Grading and Feedback: AI can automate the grading process for structured assessments and provide instant, detailed feedback on student work, identifying areas for improvement. AI can also generate **multiple-choice questions, short-answer prompts, or essay topics** based on course material.

Formative Feedback: AI is highly valuable in providing **formative feedback** to doctoral students and others, which is timely and personalized. Students should use AI to **provide feedback and suggest stronger language** for their writing.

Comparative Evaluation and Self-Assessment: Students should be encouraged to use AI for self-assessment and **comparative evaluation**, comparing their work or peer feedback against AI-generated suggestions. This practice reinforces the importance of reading critically and evaluating the AI's output.

Human Oversight in Grading: While AI can assist in grading, especially for structured assignments, its reliance on past data may penalize unconventional thinking. Instructors **must review how the AI assessed student work** and ensure that the suggestions align with academic expectations and do not introduce biases.



Regardless of the specific methodology, the success of AI integration depends on cultivating **AI literacy**.

Training and Competency: Academic institutions must provide **adequate training and adaptation periods** for researchers and educators to become proficient in utilizing AI tools effectively and ethically. This is crucial for equipping faculty to integrate AI into pedagogically sound, ethical, and **discipline-specific frameworks**.

Ethical and Responsible Use: Students must be taught to use GenAI **ethically, responsibly, and transparently**. Policies must provide clear and nuanced guidelines, rather than remaining silent, as the conversation must move beyond acceptance versus rejection.

Critical Inquiry and Stewardship: Teaching ethical AI literacy should establish that AI use requires **critical inquiry** rather than simply following rules. Training should focus on developing skills in **information verification, response integration, and task stewardship**, where the human user retains the responsibility and accountability for the work, even if the production labour is delegated to the GenAI tool.

Universal Design for Learning (UDL): AI should also be leveraged to support **social justice in higher education** by supporting the Universal Design for

Learning (UDL) approach, which helps break down barriers for individuals with disabilities or language barriers by offering personalized and adaptive experiences.

Analogy: Integrating AI tools into a Computer Science curriculum is like providing every student with a state-of-the-art computational laboratory and a patient, expert research assistant (the AI). The effective methodology requires the professor to stop grading the lab manual entry (the final product) and start grading the student's ability to critically design experiments, interpret the assistant's results, and make ethical, accountable decisions—thereby focusing on the *intelligence* and *judgment* that the human uniquely provides.

ACKNOWLEDGEMENTS

https://notebooklm.google.com/notebook/313f3876-d72e-4855-bbb1-df4b44174ab7

I used AI in the following ways:

(i) I have previously created a relevant PPT for Master students

I used traditional research methods and social media to collect resources.

I used social media aggregation: LinkedIn(https://www.linkedin.com/) and X (https://x.com/) to collect relevant academic literature and infographics. I further edited the text adding appropriate information and created the PPT myself.

(ii) I generated further ideas or structure suggestions and preparatory activity for the assignment,

I used Google Scholar (https://scholar.google.com/) and Perplexity (https://www.perplexity.ai/) to find relevant academic literature and my own LinkedIn(https://www.linkedin.com/) to be offered relevant literature according to my interest. I used Microsoft Word with OpenAI with ChatGPT5.1 (https://chatgpt.com/) to prepare the structure (20 iterations).

(iii) I created the draft text of each chapter using appropriate prompts to extract information from the collected set of 41 scholarly articles to suit the chapters of the assignment,

I used NotebookLM (https://notebooklm.google/ to compose the main texts of the chapters.

(iv) generated text, rewrite, rephrase and/or paraphrase a portion of this assignment.

I used Microsoft Word with OpenAI with ChatGPT5.1 (https://chatgpt.com/) to prepare the drafts (20 iterations). I further edited the text adding appropriate information.

(v) I created the PPT.

I tried Beautiful.ai to make the PPT but found unsuitable for a concise and appropriate message. I edited the final PPT using manually extracted text for the slides, adding appropriate information from the Word file into the Notes.

Legend: [1] AI system(s); [2] link to the tool; [3] how the tool was used or prompts included; [4] edits and changes made to the output.