Workshop Outline: Who Owns the Output? Fair Use and IP in Generative AI

Slide 1: Introduction: The Promise, The Catch, and The Law

- **Time:** 5 minutes
- **Focus:** Framing the current landscape objectively.
- **Talking Points:**
    - Generative AI disrupts traditional notions of intellectual property on two distinct fronts: how the models are built (inputs) and what the models produce (outputs).
    - **Goal:** Demystify the current legal realities established by the U.S. Copyright Office and federal courts so faculty can innovate confidently and understand the rights attached to their scholarship.
- [Placeholder: Insert Your Learning Objectives Here]
- **[Placeholder: Insert Your Script Here]**

Slide 2: The Government's Stance: Overview of the U.S. Copyright Office Reports

- **Time:** 10 minutes
- **Focus:** Establishing the baseline rules as currently understood by the U.S. government.
- **Talking Points:**
    - **Part 1 (Digital Replicas):** Examines the unauthorized replication of voice or visual likenesses, highlighting the risks of deepfakes and the legislative gap in federal right-of-publicity protections.
    - **Part 2 (Copyrightability):** Addresses the output problem, reaffirming that copyright protection requires a human author and that purely AI-generated material cannot be registered.
    - **Part 3 (Generative AI Training):** Addresses the input problem, focusing on the massive ingestion of copyrighted works to train AI models and the complex debate over whether this constitutes infringement or fair use.
- **[Placeholder: Insert Your Script Here]**

Slide 3: The Input Problem: Training AI & The Fair Use Doctrine (Report Part 3)

- **Time:** 10 minutes
- **Focus:** How AI models learn and the controversy over using copyrighted academic works.
- **Talking Points:**
    - **The Ingestion Process:** To train generative AI models, tech companies mass-download datasets, which inherently involves making copies of millions of copyrighted works.
    - **The Fair Use Defense:** Technology companies rely on the Fair Use doctrine, arguing their ingestion of these works is highly "transformative" because it creates a new tool rather than competing with the original text.
    - **The Market Harm Argument:** Authors and publishers focus on the market effect, arguing this mass ingestion threatens their livelihoods, replaces the need for original works, and should require licensing.
    - **Discussion Prompt:** If a commercial AI model ingests decades of published, peer-reviewed journal articles to train its text-generation capabilities without compensating the original authors or publishers, how does this impact the academic publishing ecosystem and the concept of fair use?
- **[Placeholder: Insert Your Script Here]**

Slide 4: The Baseline Rule of Authorship (*Thaler v. Perlmutter*)

- **Time:** 10 minutes
- **Focus:** The strict legal precedent that machines cannot be authors.
- **Talking Points:**
    - Dr. Stephen Thaler attempted to register an artwork, explicitly listing his AI system (the "Creativity Machine") as the sole author.
    - The U.S. Court of Appeals affirmed that the Copyright Act of 1976 requires all eligible work to be authored in the first instance by a human being.
    - **The Rationale:** Because machines lack the legal capacity to hold property, experience a human lifespan, or form intent, they cannot be granted copyright monopolies.
    - **The Precedent:** This establishes the absolute floor—zero human intervention means zero copyright protection.
- **[Placeholder: Insert Your Script Here]**

Slide 5: The Boundary of Human Control (*Allen v. Perlmutter* & *Feist*)

- **Time:** 15 minutes
- **Focus:** Comparing *Allen* to *Thaler* to explore the inconsistencies in how the authorship test is applied, grounded in the threshold for originality.
- **Talking Points:**
    - **The Originality Threshold (*Feist v. Rural Telephone*):** The Supreme Court established that the bar for copyright originality is "extremely low," requiring only "independent creation plus a modicum of creativity".
    - **The Contrast:** Unlike Thaler, who claimed no human authorship, Jason Allen argues he is the author of an AI-generated image because he engaged in an iterative process of over 624 prompts to shape the tone, composition, and style.
    - **Allen's Argument:** Allen argues his prompting easily meets the *Feist* threshold, and the AI was merely an "assisting instrument"—much like a camera in early photography copyright cases—translating his specific creative vision into tangible expression.
    - **The Copyright Office's Rejection:** The Office argues that prompts amount only to unprotectable "ideas," and that the AI executes the actual "expression," conflating the idea/expression dichotomy.
    - **The Fairness Debate:** Allen's brief argues the Office is imposing an impossible "sufficient control" standard that selectively penalizes the randomness in AI tools, while accepting far greater randomness in tools like photography.
    - **Discussion Prompt:** The Supreme Court says the bar for creativity is "extremely low." If a researcher spends hours engineering highly specific prompts to generate a complex data visualization, are they just providing an "idea," or are they executing creative control? Is the Office holding AI to a higher standard than other technological tools?
- **[Placeholder: Insert Your Script Here]**

Slide 6: Practical Strategies & Open Q&A

- **Time:** 10 minutes
- **Focus:** Actionable takeaways for faculty teaching, research, and publishing.
- **Talking Points:**

    - **The Disclaimer Rule:** Under current Copyright Office guidance, applicants must explicitly disclose and disclaim any AI-generated content that is more than *de minimis* when registering a work.
    - **Journal Policies:** Most academic publishers have adopted strict guidelines requiring explicit disclosure of AI use in the research or writing process.
    - **Protecting Unfinished Work:** Faculty should be cautious about inputting their own unpublished research data into public AI tools to avoid inadvertently feeding future training models.
- **[Placeholder: Insert Your Script Here]**