quinn emanuel trial lawyers

quinn emanuel urquhart & sullivan, llp

POSITION & PROTECT:
THE LEGAL STATE OF PLAY IN AI

AI in Finance Summit
April 19, 2024

THE IMF PREDICTS:

"FINANCIAL INSTITUTIONS ARE FORECAST TO DOUBLE THEIR SPENDING ON AI BY 2027"
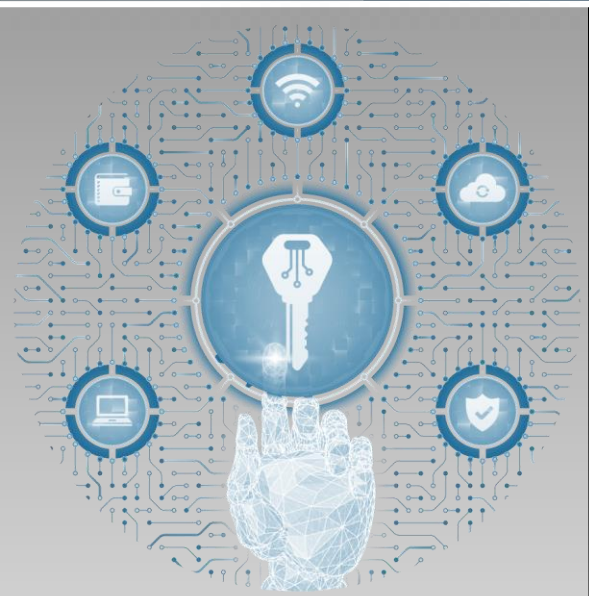
AGENDA

1 Training AI

2 Protecting AI

3 Regulatory Considerations

4 Best Practices

quinn emanuel trial lawyers
quinn emanuel urquhart & sullivan, llp

1 **TRAINING AI**
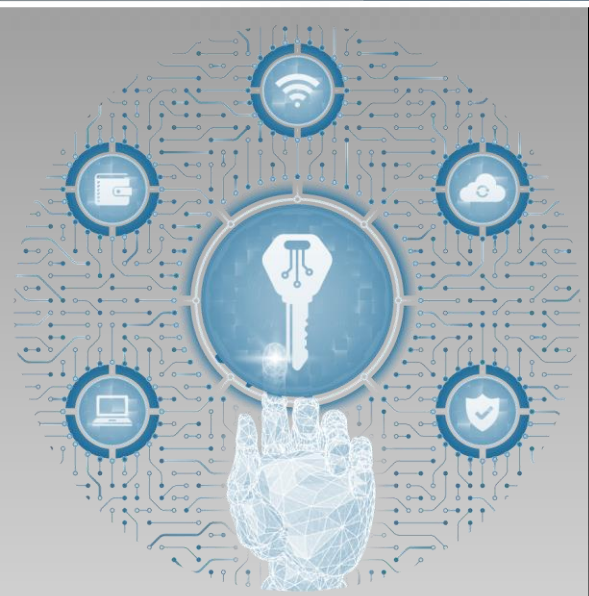
TRAINING AI: WEB SCRAPING

- Sometimes called data extraction or aggregation, web scraping is the practice of collecting vast amounts of online data by automated means.

- Web scraping (or purchasing scraped data from a vendor) is one common method of obtaining training data.

Is web scraping legal?

TRAINING AI: SCRAPING & THE COMPUTER FRAUD AND ABUSE ACT (CFAA)

- CFAA imposes both civil and criminal liability for computer hacking.

- CFAA prohibits accessing a computer without authorization.

- Scraping data behind a login screen or paywall could be a CFAA violation if the scraper is **not authorized** to access the data.

- CFAA does **not** apply to **publicly-available data.**

"Giving companies like LinkedIn free rein to decide, on any basis, who can collect and use data—data that the companies do not own, that they otherwise make publicly available to viewers, and that the companies themselves collect and use— *risks the possible creation of information monopolies that would disserve the public interest.*" *hiQ v. LinkedIn*

- Many website's Terms of Service (TOS), on their face, prohibit data scraping.

- Many website's TOS, on their face, also prohibit commercial use of data.

- TOS may constitute a binding and enforceable contract.

- **Background:** Bright Data sells data scraped from Meta's Facebook and Instagram platforms.

  - Data is scraped in a "logged-off" state – no account logged in at the time of collection.

  - Meta sued, claiming this scraping violates Meta's TOS.

  - In January 2024, the California Federal Court ruled that the TOS did <u>not</u> prohibit "logged-off" scraping.

- **Key Findings:**

  - **TOU – Logged-On "Use" Only**: Meta's TOU held to apply only to account holders while they are logged in.

    - Users were not bound by TOU restrictions for logged-off usage, and "those who do not subscribe to Meta services do not see the Terms and cannot be bound thereby."

  - **Perpetual Restrictions Unenforceable**: "Survival clause" in Meta TOU purporting to prohibit scraping in perpetuity—even after a user terminated any agreements with Meta—was not enforceable.

  - **CAPTCHA Not a Login**: A CAPTCHA designed to deter scraping does not convert public information to "private" the same way as a log-in screen.
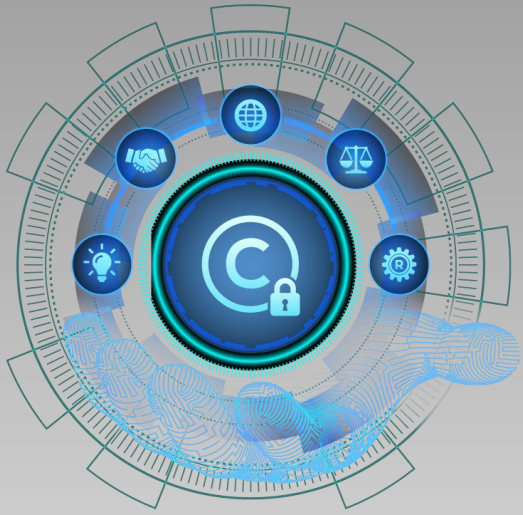
- **Unauthorized use of copyrighted work to train AI may infringe copyrights.**

- **But this is unsettled territory:**

  - **Use of data during training:** Is an unauthorized "copy" created when the model processes training data as input? Is using the copyrighted work for training a "fair use" or "transformative"?

  - **Data stored in the model:** Does the model itself include a copy of copyrighted material? Can the weights and parameters in the model be considered "derivative" works?

  - **Outputs of the model:** Do the outputs of the model reproduce copyrighted works? If the model generates outputs in the "style" of an author or artist, is that a "derivative" work?

# AI-GENERATED CONTENT: LICENSING AND RIGHTS OF PUBLICITY
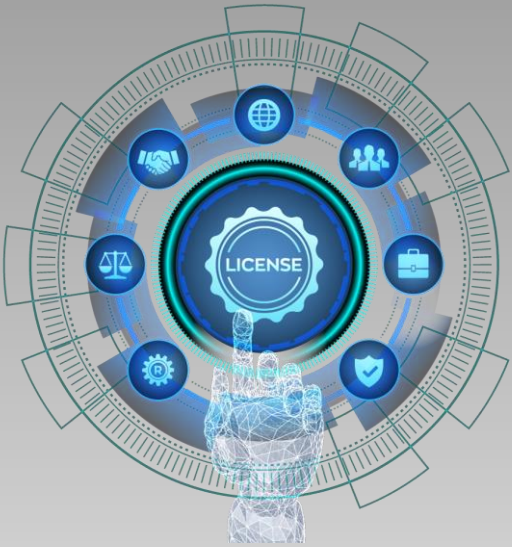
- Generating content from a model trained on licensed content could trigger licensing terms

  - E.g., AI-generated source code or images trained on open source or creative commons material could trigger attribution/disclosure terms under those licenses

- AI-generated content of real people may have "right of publicity" implications

  - E.g., use of AI tools to generate speech using the voice of a musician or voice actor, or use of AI tools to insert a celebrity into an image or video

quinn emanuel trial lawyers
quinn emanuel urquhart & sullivan, llp

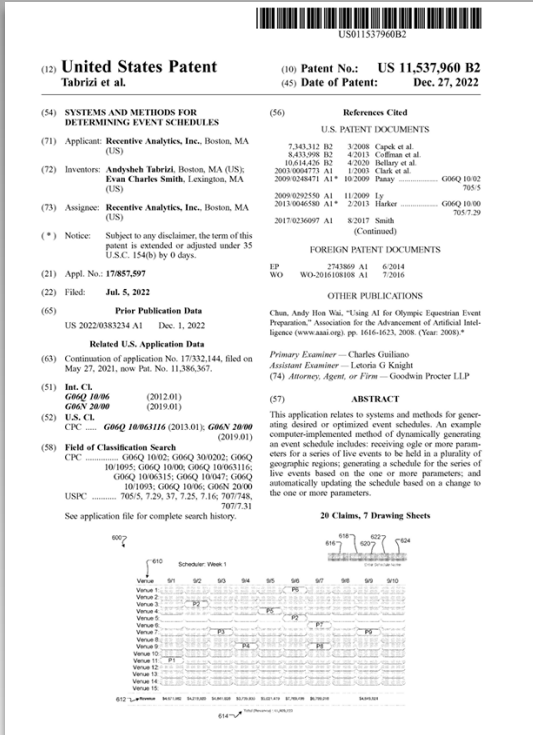**2** PROTECTING YOUR INTELLECTUAL PROPERTY FOR AI

COPYRIGHT AND AI-GENERATED CONTENT
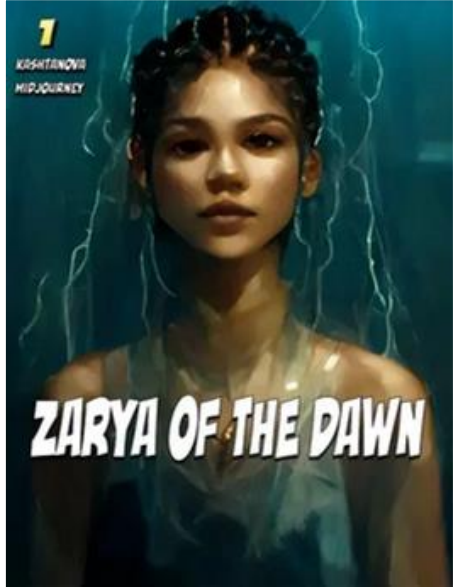
- AI-generated inventions not entitled to patent protection under U.S. law (only humans can be inventors)

- But human arrangement or modification of AI-generated content could be protectable.

- A creative AI prompt may be copyrightable, but the AI-generated product of the prompt is not

- Software could be copyrightable even if portions compromise AI-generated code

- **Completely AI-generated content not entitled to copyright protection under U.S. law (only human authors)**

- **Inventions covering AI technology may be difficult to patent**
  - **Many AI patents invalidated by courts as unpatentable subject matter**

- **AI inventions that are patentable may provide competitors a road map to success**

- **Trade secrets may be a more effective route for protecting AI inventions**

- **Pros:**

  - No limitation on subject matter or inventorship

  - Protects against violations outside the U.S.

  - Protected indefinitely, as long as it is kept secret

  - Harder to design around a trade secret

- **Cons:**

  - No protection against reverse engineering/independent development

  - Protection may be forfeited if "reasonable measures" not taken

3 AI REGULATORY TRENDS

AI REGULATORY TRENDS: EU AI ACT

Bans Certain Practices

- Biometric categorization and untargeted scraping of facial images

- Emotion recognition at work and schools
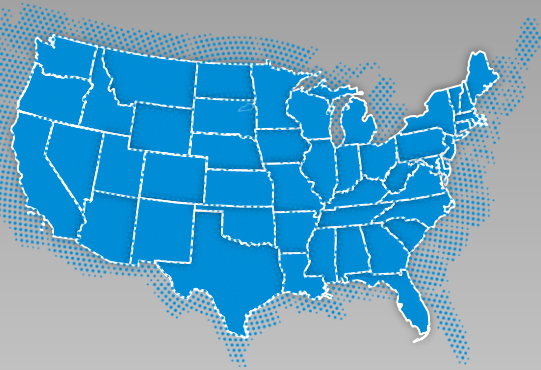
- Social scoring

Requires Transparency

- Requires labeling of AI-generated content

- Essential services must conduct impact assessment

- For larger AIs, disclosure of training data, security procedures, and energy efficiency

New Regulators

- European AI Office to enforce laws

- **California Automated Decision Tools Bill ("AB 331"):**
  - Requires impact assessments and disclosure of training information for "Automated Decision Tools"
    - "Automated Decision Tools" – systems or services that use AI to make decisions which have a legal, material, or other significant effect on an individual's life in terms of the impact of, access to, or cost, terms, or availability of employment, education, housing, healthcare, financial services, and more.

- **California, Colorado, Connecticut, Virginia, Utah, Tennessee, Montana, Indiana, Delaware Expand Data Privacy Protections**
  - Grant consumers rights to opt-out of profiling for automated decisions and requires data protection assessment for activities that pose a "heightened risk of harm," including targeted advertising and some types of profiling.

- **No overarching national data privacy law in US**

4 **BEST PRACTICES FOR AI INTEGRATION**

MANAGING RISK: DATA SECURITY IN CHAT BOTS

- **Ensure that prompts and output are not being saved or used to train the AI model.**
  - ChatGPT and other tools have this as an option in "settings."

- **Considerations for any Generative AI tools—public or private.**
  - What is the model of AI being used - open, closed, public versions, private versions, upgraded or enterprise public versions?
  - Is input data anonymized? If so, to what anonymization standard?
  - Is the AI tool able to provide its own dedicated tenant, so data resides within your organization's control? (I.e., Microsoft Copilot does this for security).
  - Can each party in the data flow produce audit logs of user requests?
  - Are search results retained according to your organization's retention requirements to review activity as needed?
  - Can parties throughout the data flow request and review security audit and penetration test results of the AI tools to identify any vulnerabilities?

- Generative AI is prone to "hallucinations" – *i.e.*, fabricating facts, authority, or other output.

- AI tools may summarize material inaccurately, change language, or draw on materials from outside of the intended research pool.

- Bias can carry over from training sets.

- Inherent tradeoff between "creativity" and hallucination.

quinn emanuel trial lawyers
quinn emanuel urquhart & sullivan, llp

- Only ask about public information.

- Human review of any output is necessary.

  - GenAI is an assistant *not* a replacement.

- Review the privacy policies and terms of services *prior* to use.

- Identify any work product generated by GenAI.

- Know the applicable rules and policies.

- Create an in-house AI Policy with weigh-in from legal, compliance, and business teams.

quinn emanuel trial lawyers
quinn emanuel urquhart & sullivan, llp

quinn emanuel trial lawyers
quinn emanuel urquhart & sullivan, llp

QUESTIONS?

hopeskibitsky@quinnemanuel.com

ronhagiz@quinnemanuel.com