

# NCI MIDI Workshop

Advances in Medical Imaging  
De-Identification and the Impact of  
Regulatory Constraints

**Abraham Gutman**  
President & CEO, AG Mednet

# State of the Art in Pixel De-identification of DICOM Images

- Tool to delineate PHI by region and apply pixel removal across multiple images
- Can be problematic if series are misaligned
- May require individual region redaction by user
- Two possible next gen approaches in prototype stage
  - Define modality “formatting” and automatically remove region
  - Use ML and LLMs to find text in images and remove names
    - Have prototyped using AWS Rekognition
    - Requires uploading images containing PHI!



# Harnessing AI and LLMs in De-identification of DICOM Images

- Leverage trained model for both pixels and structured reports
- How do AI solutions operate? Computer-assisted or autonomous
- Scalability and adaptability of AI solutions to handle large datasets and various medical imaging formats: Where will this run?
- If de-identifying video, will processing time be an issue?



# Overcoming Regulatory Hurdles in De-identification of DICOM Images

- Exploring regulatory changes to allow de-identification in the cloud
- Will investigator sites allow local software installation?  
(Is that practical?)

The technology is here to enable ML/LLM-driven de-id.

Investigator sites' IT departments are not ready



# Navigating Machine Learning Integration in GxP Compliant Systems

- Addressing the dynamic nature of machine learning models in the context of GxP compliance requirements
- Establishing validation processes for AI and LLM-driven de-identification tools within GxP frameworks
- What does Change Control, Traceability and Reproducibility mean for machine learning models?

Need to collaborate with regulatory authorities to develop clear guidance for integrating machine learning systems in GxP environments

