Public Libraries
& AI

THE GALECIA GROUP
technology solutions for librarians

LIBRARY OF MICHIGAN

1. Transparency: The "Black Box" Problem

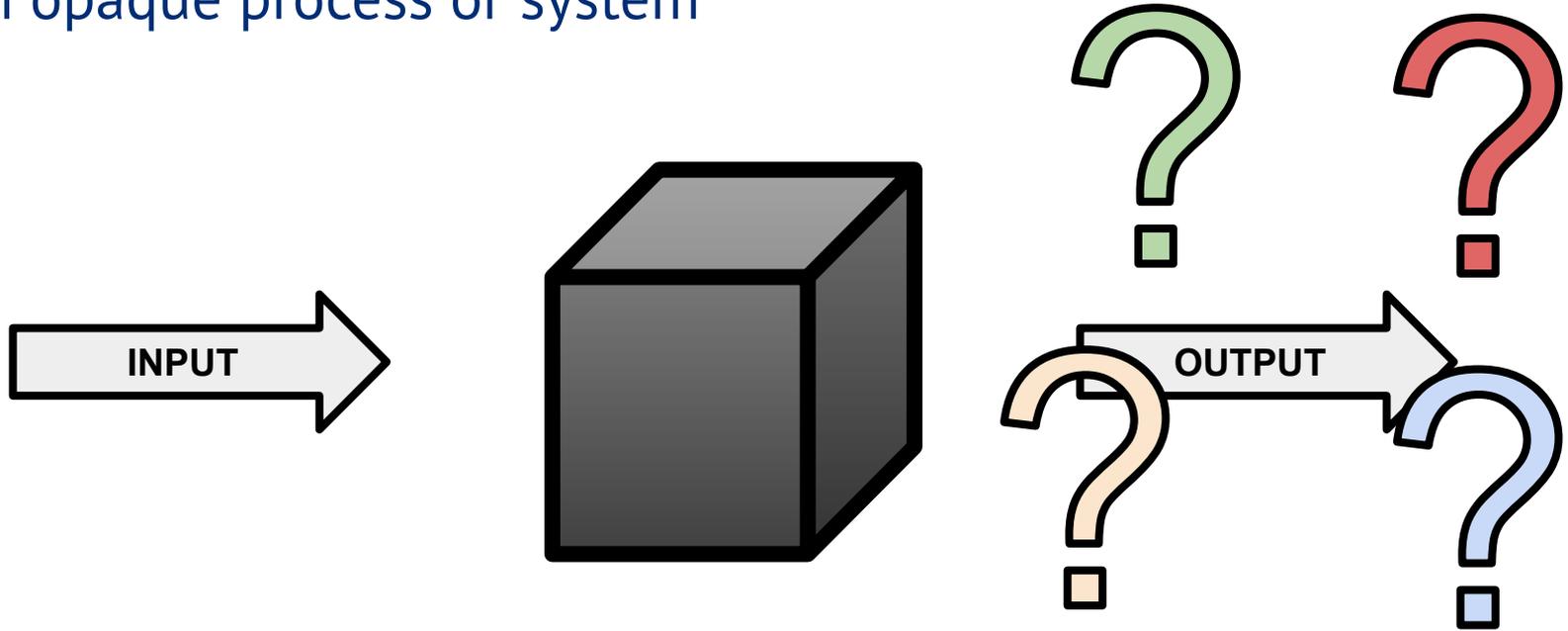Series 2: AI Risks and Limitations

The "Black Box" Problem

- An opaque process or system

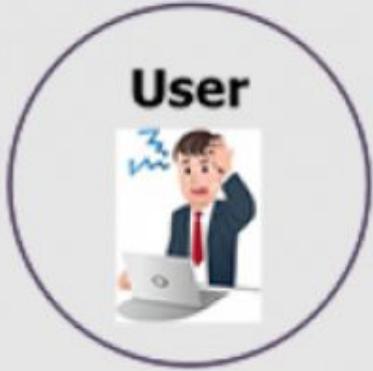# The "Black Box" Problem

- An opaque process or system

What is the AI "learning" from?

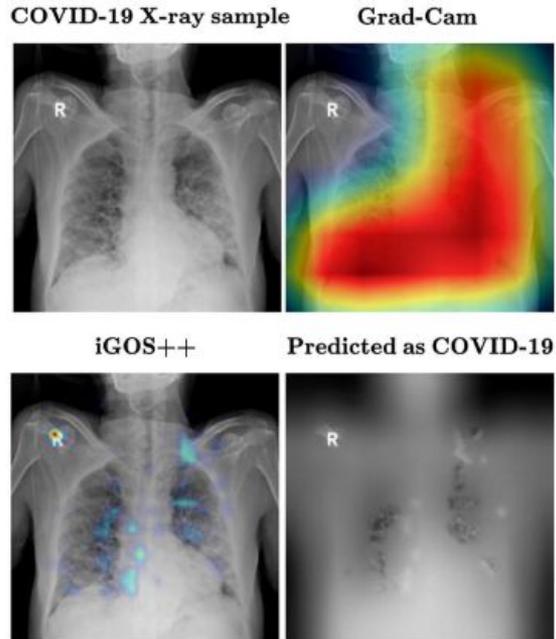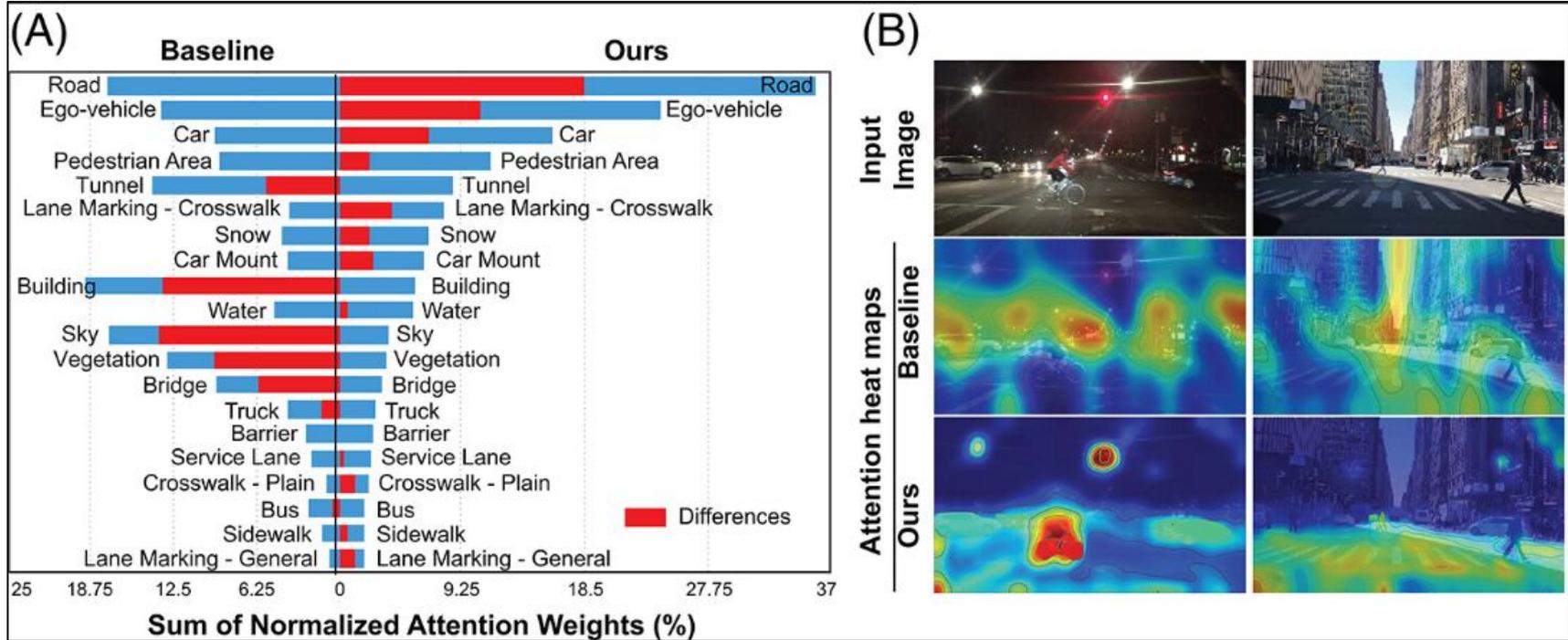| Source	Description
CommonCrawl	Public Web Crawler Data
Wikipedia	Online Encyclopedia
Reddit	Community Forums
PubMed	Medical / Biology Archive
GitHub public repos	Open Source Software Code
Project Gutenberg	Public Domain Book Archive
"Books1, Books2, Books3"	Pirate / Shadow Libraries

The "Black Box" Problem

Why It Matters

COVID-19 X-ray sample   Grad-Cam

iGOS++   Predicted as COVID-19

Li, F., Qi, Z., Khorram, S., Shitole, V., Tadepalli, P., Kahng, M., & Fern, A. (2021). From heatmaps to structured explanations of image classifiers. *Applied AI Letters*. https://doi.org/10.1002/ail2.46
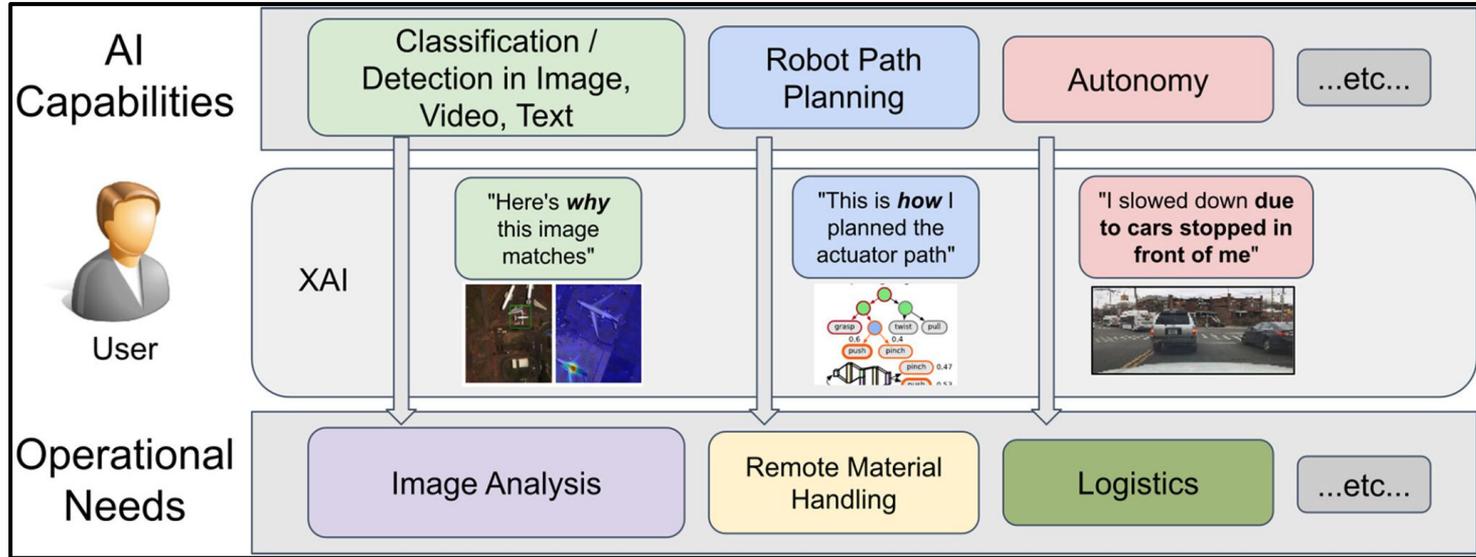
Why It Matters

Kim, J., Rohrbach, A., Akata, Z., Moon, S., Misu, T., Chen, Y.-T., Darrell, T., & Canny, J. (2021). Toward explainable and advisable model for self-driving cars. *Applied AI Letters*. https://doi.org/10.1002/ail2.56

Why It Matters

Explainable AI ("XAI")

- Helps humans "see" inside the AI processes

Hu, B., Tunison, P., Vasu, B., Menon, N., Collins, R., & Hoogs, A. (2021). XAITK: The explainable AI toolkit. *Applied AI Letters*. https://doi.org/10.1002/ail2.40

Coming up...

- More risks and limitations: accuracy, context, and common sense

Thanks for watching!