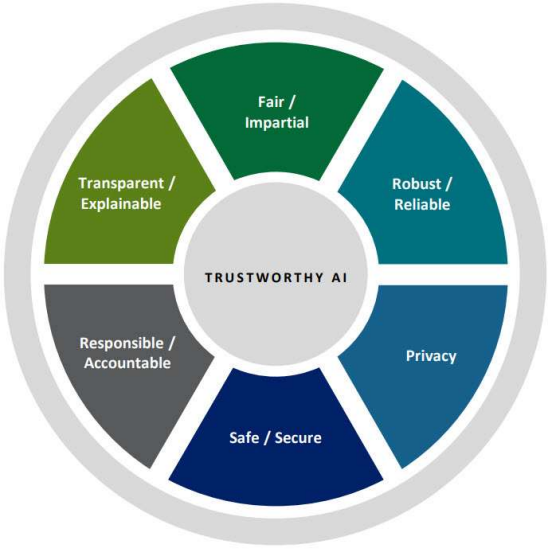
Short summary of Trustworthy AI Principles



Fair / Impartial

- Ensure equitable application across all participants
- Validate input data and AI output across sub-populations

Robust / Reliable

- Produce accurate, consistent, and reliable outputs
- Be monitored and updated to reduce errors

Transparent / Explainable

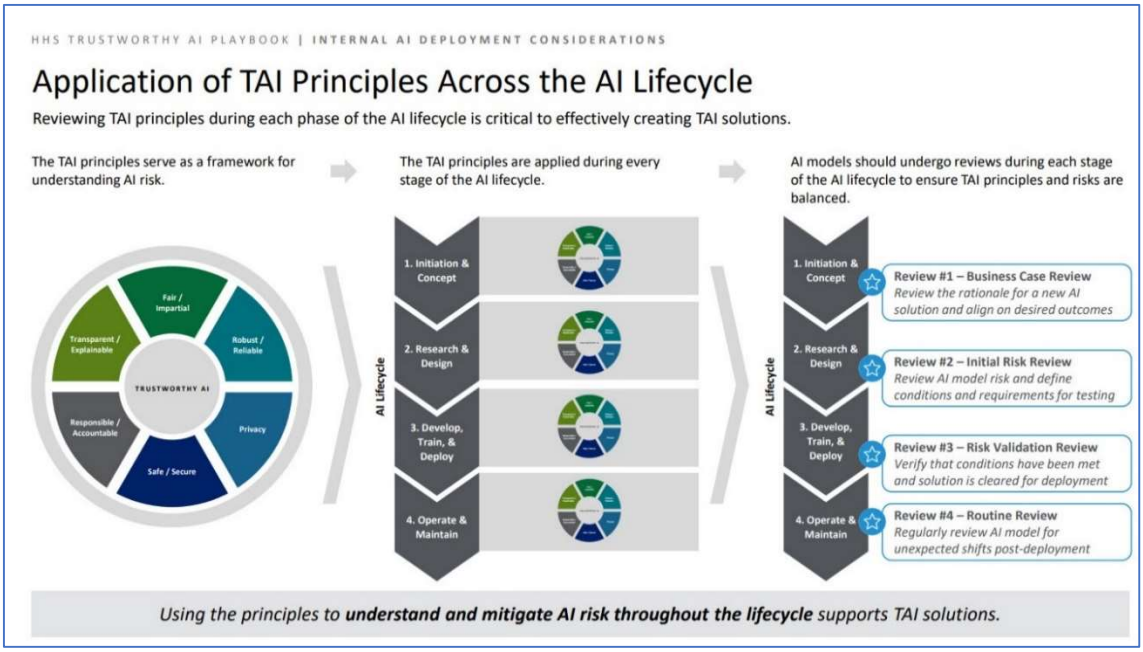
- Explain how input data is being used
- Make clear which factors drove systems to make decisions
- Open up algorithms, attributes, and correlations for inspection

Privacy

- Respect individual, group, or entity privacy
- Confine data to its intended and stated use
- Ensure data used has been approved by the data owner or steward

Safe / Secure

- AI systems should be protected from risks (including Cyber) that may directly or indirectly cause physical and/or digital harm to any individual, group, or entity (incl. HIPAA, legal & clinical risks)



Source: hhs.gov/sites/default/files/hhs-trustworthy-ai-playbook.pdf
UCSF reference: https://ai.ucsf.edu/trustworthy
Dated: 1/22/24