Summary of *Who Cares About Trust? Clusters of Research on Trustworthy AI*

This report identifies **18 clusters of research papers in CSET's Map of Science that are relevant to the development of trustworthy artificial intelligence**, along with the most referenced papers among those within the clusters. From these clusters, the authors find:

- **The 18 trustworthy AI clusters cover a broad spectrum of AI methods, techniques, and applications,** including deep learning, adversarial attacks, word embeddings, image privacy, speech recognition, explainable AI, federated learning, algorithmic fairness, differential privacy, and robotics. (See figure below.)
- Given their wide distribution across AI research clusters, the terms ***reliability***, ***safety***, **and** ***robustness*** **appear to be widely studied and/or adopted metrics for AI systems**. By contrast, ***interpretability***, ***transparency***, ***explainability***, ***security***, **and** ***privacy*** **appear to be specific issue areas,** as they are concentrated in fewer clusters.

Cluster Key Concept Groups and Trustworthy AI Keyword Term Appearance

✓ indicates 10% or more of the publications in the clusters included the keyword

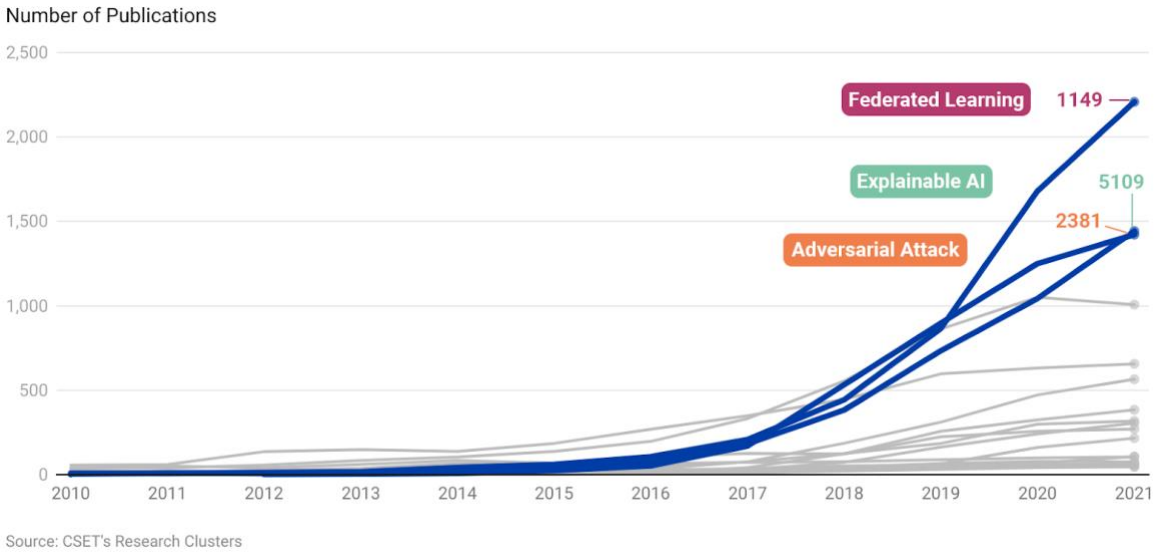
	Bias	Explainability	Fairness	Interpretability	Privacy	Safety	Security	Reliable	Robustness	Trust
Adversarial Attacks							✓		✓	
Algorithmic Fairness	✓		✓							
Deep Learning		✓		✓		✓		✓	✓	✓
Differential Privacy					✓					
Explainable AI		✓								✓
Federated Learning					✓					
Image Privacy					✓					
Robotics Security							✓			
Speech Data					✓					
Word Embeddings	✓									

Source: CSET's Research Clusters

AI Research Cluster Key Concepts and Trustworthy AI Keyword Term Appearance

- **Three research clusters stood out: one related to adversarial attacks (2381), one related to federated learning (1149), and one related to explainable AI (5109).** All three experienced rapid growth since 2017 and, at the same time, the percentage of papers using a trustworthy AI keyword within the cluster also grew. (See figure below.)

Growth of Trustworthy AI Keyword Clusters Over Time, 2010-2021



Source: CSET's Research Clusters

For more information:

- Download the report: https://cset.georgetown.edu/publication/who-cares-about-trust
- Contact: Autumn Toney (Autumn.Toney@georgetown.edu).