Сбер разработал и опубликовал первую в России комплексную модель киберугроз, связанных с системами искусственного интеллекта (ИИ). Документ охватывает все этапы жизненного цикла таких систем — от подготовки данных до внедрения готовых решений в приложения. Материалы доступны на портале «Кибрарий».

Созданная модель включает 70 различных угроз, которые могут возникнуть при использовании генеративных и предиктивных ИИ-моделей. Каждая угроза описана с точки зрения потенциальных последствий, затрагиваемых информационных свойств (например, конфиденциальности или целостности), а также возможных целей атаки — таких как обучающие датасеты или открытые модели.

Модель призвана помочь компаниям из любых отраслей выявлять уязвимости в своих ИИ-системах, выстраивать подходы к их защите и снижать возможные риски. Она содержит наглядные схемы и пояснения по взаимодействию компонентов ИИ, что упрощает применение документа на практике.

Сбер создавал модель, опираясь на собственный опыт в обеспечении безопасности своих ИИ-разработок, а также на международные практики, включая материалы OWASP, MITRE и NIST.