17.06.2021

Kurz-Glossar

Einige zentrale Begriffe aus dem Bereich Medizin-KI & Datenschutz

- **Anonymisierung und Pseudonymisierung**
  Ansätze, um den Datenschutz in Datensätzen zu gewährleisten. Bei der Pseudonymisierung werden identifizierende Daten wie Name oder Anschrift durch ein Pseudonym ersetzt. Bei der Anonymisierung werden alle Merkmale und Daten, die auf eine Person bezogen oder beziehbar sind, gelöscht, aggregiert, getrennt oder verfälscht, sodass die einzelnen Datenpunkte untereinander nicht mehr verknüpfbar sind. Das Ziel ist, die Zuordnung der Daten zu beispielsweise einer Person auf diese Weise unmöglich zu machen.
- **Deanonymisierung**
  Gezieltes Aufheben der Anonymisierung. Durch Kombination mit anderen Datensätzen, weiteren Informationen oder unzureichende Anonymisierung können eigentlich anonymisierte Daten wieder auf beispielsweise einzelne Personen bezogen werden.
- **Federated Learning / Föderales Lernen**
  Eine Methode zum Anlernen eines KI-Modells, bei der die zum Trainieren verwendeten Daten bei den „Besitzern" blieben. Eine zentrale Instanz stellt den anzulernenden KI-Algorithmus bereit, dieser wird lokal da angelernt, wo die Daten bereits liegen. Parameter der so angelernten lokalen KI-Modelle werden an die zentrale Instanz zurückgesendet, wo sie zur Aktualisierung des gemeinschaftlichen KI-Modells verwendet werden. Aktualisierte Parameter werden dann an die lokalen Instanzen zurückgeschickt und das dortige Modell wird weitertrainiert. Der Vorgang kann wiederholt werden, bis eine gewisse Anzahl Wiederholungen oder ein Schwellenwert in der Performance des Modells erreicht wurde. So erhält die zentrale Instanz nie Zugriff auf die zum Trainieren verwendeten Daten selbst.
- **Model Inversion Attack**
  Eine Methode, bei der aus den Parametern eines angelernten Modells Rückschlüsse auf die Trainingsdaten gezogen werden. Das reicht von Sprachmodellen, die eingegebene Sätze mit Angaben aus den Trainingsdaten vervollständigen, bis zur Rekonstruktion von (wenn auch unscharfen) Bildern aus den Trainingsdaten.
- **Reconstruction Attack**
  Eine Methode, durch die auf nicht öffentliche Teile eines Datensatzes geschlossen werden kann. Dabei wird auf einen öffentlichen aggregierten Datensatz oder Statistiken über diesen Datensatz zugegriffen, wobei Rückschlüsse auf private Elemente des Datensatzes oder eine teilweise Rekonstruktion dessen möglich ist.
- **DSGVO**
  Datenschutz-Grundverordnung. Eine 2016 in Kraft getretene Verordnung der EU, in der EU-weite Regeln zur Verarbeitung personenbezogener Daten festgelegt wurden.