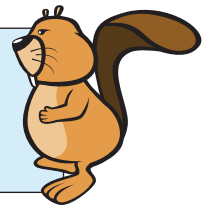


Stufen	5 – 6	leicht	mittel	schwer
Stufen	7 – 8	leicht	mittel	schwer
Stufen	9 – 10	leicht	mittel	schwer
Stufen	11 – 13	leicht	mittel	schwer



De-Anonymisierung

Krankenakten enthalten sensible persönliche Daten, die nicht öffentlich bekannt sein sollten. Zu Forschungszwecken hat ein Krankenhaus aktuelle Daten deshalb anonymisiert veröffentlicht; die Tabelle zeigt einen Auszug aus dieser Liste:

Geb.datum	Geschlecht	PLZ	Krankheit
01.01.1974	männlich	29400	Diabetes
01.01.1976	männlich	18250	Lungenkrebs
01.01.1976	weiblich	29400	Brustkrebs
01.01.1976	weiblich	29400	Fehlgeburt
01.01.1984	weiblich	18250	Herzanfall
01.01.1985	weiblich	16300	Brustkrebs
01.01.1987	weiblich	25340	Hautkrebs
01.01.1988	männlich	18250	Diabetes
01.01.1988	weiblich	18250	Grippe

Gleichzeitig hat – wegen anstehender Wahlen – die Gemeinde mit PLZ 18250 eine Liste der Wahlberechtigten veröffentlicht; die Tabelle zeigt die Daten **ALLER** Wahlberechtigten, die an einem 1. Januar geboren wurden:

Geb.datum	Geschlecht	Name
01.01.1958	weiblich	Melanie Meyer
01.01.1976	männlich	Georg Schmidt
01.01.1976	männlich	Robert Schlumpf
01.01.1984	weiblich	Kathrin Frei
01.01.1984	weiblich	Eva Müller
01.01.1988	weiblich	Agnes Bachmann
01.01.1988	männlich	Roman Schröder
01.01.1988	weiblich	Isabelle Beyer
01.01.1989	männlich	Martin Klaus

Anhand der beiden Tabellen kannst du eine Person mit Namen identifizieren (de-anonymisieren), die mit absoluter Sicherheit eine Krankheit hat.

Wie lautet der Name dieser Person?

- A) Georg Schmidt
- B) Eva Müller
- C) Roman Schröder
- D) Isabelle Beyer



Stufen	5 – 6	leicht	mittel	schwer
Stufen	7 – 8	leicht	mittel	schwer
Stufen	9 – 10	leicht	mittel	schwer
Stufen	11 – 13	leicht	mittel	schwer

Antwort C ist richtig:

Die Patienten der Zeilen 1, 3, 4, 6 und 7 können es nicht sein, da sie nicht in der Gemeinde mit der Postleitzahl 18250 wohnen.

Der Patient der Zeile 2 ist im Jahr 1976 geboren, männlich und wohnt in der Gemeinde mit der Postleitzahl 18250. Es gibt aber zwei Einwohner, die diesen Daten entsprechen: Georg Schmidt und Robert Schlumpf.

Die Patientin der Zeile 5 ist im Jahr 1984 geboren, weiblich und wohnt in der Gemeinde mit der Postleitzahl 18250. Es gibt aber zwei Einwohnerinnen, die diesen Daten entsprechen: Kathrin Frei und Eva Müller.

Die Patientin in der Zeile 9 ist im Jahr 1988 geboren, weiblich und wohnt in der Gemeinde mit der Postleitzahl 1825. Es gibt aber zwei Einwohnerinnen, die diesen Daten entsprechen: Agnes Bachmann und Isabelle Beyer.

Der Patient in der Zeile 8 jedoch, geboren im Jahr 1988, männlich und in der Gemeinde mit der Postleitzahl 1825 wohnend, lässt sich eindeutig als Roman Schröder identifizieren.

Das ist Informatik!

Wenn Daten über einen Menschen gespeichert sind, dann kann man diesen Menschen anhand seiner Daten eindeutig wieder erkennen. Oder nicht: Dann bleibt der Mensch anonym.

Ob jemand anonym bleibt, hängt von der Bedeutung der gespeicherten Daten und von der Anzahl der potentiell in Frage kommenden Menschen ab. Ist zum Beispiel über jemand nur der Name "Müller" und das Gewicht "ca. 85 Kilo" bekannt, dann bliebe er in einer Millionenstadt sicher anonym – da gibt es viele Müllers dieser Gewichtsklasse – in einem Mietshaus mit nur vier Parteien wohl eher nicht.

Legt man für verschiedene Zwecke unabhängig voneinander zustande gekommene Datensammlungen zusammen und alle Betroffenen waren hier wie da anonym, dann kommt diese Anonymität in Gefahr. Die Daten pro Mensch werden mehr und selektiver, die Anzahl der Selektierten sinkt.

Das Gegenmittel wäre die ausnahmslose Verpflichtung zur Zweckgebundenheit von Datensammlungen. Das Zusammenlegen mit anderen Daten wäre dann ein gegenüber dem ursprünglichen Sammelzweck neuer Zweck und also grundsätzlich verboten.