

## **Probleme mit Daten - Nichts Neues unter der Sonne**

*von Toni Saller*

'Big Data' begann in Mesopotamien, vor 5000 Jahren. So gelernt an der Senioren-Uni an einem Vortrag von Juraj Hromkovic über die Geschichte der Informatik. Wir erfahren, dass tatsächlich die Schrift nicht erfunden wurde, um Gedichte zu schreiben, sondern um festzuhalten, wer von den Bürgern im Reich brav die Steuern bezahlt hat. Das Reich wurde zu gross, die Beamten überfordert, um alle möglichen Steuersünder nur durch die Leistung ihrer Gedächtnisse ausfindig zu machen. Da auch die Schrift Informationen als Folge von Symbolen festhält, darf man ruhig sagen, dass Mesopotamien die erste digitale Datensammlung erstellte, die erste Datenspeicherung ausserhalb des menschlichen Hirns. Damit stellte sich bald schon das Problem der Datensicherheit. Vor 3500 Jahren wurde dafür die erste Geheimschrift entwickelt, ein Code, der die Anweisung enthielt, wie in einer verschlüsselten Schrift die Buchstaben vertauscht werden. Der Code selber durfte nicht notiert werden, um ihn vor Diebstahl zu schützen, blieb er nur im Kopf der Beamten. Codes und Geheimschrift waren dann in der Antike in vielen Gesellschaften verbreitet, in Griechenland, Ägypten, China und Indien zum Beispiel.

Und der Versuch, aus Daten Nutzen zur Beeinflussung von Bürgern und Konsumenten zu ziehen, ist auch keine Neuerfindung von Cambridge Analytica. Folgendes Beispiel: Sie wollen wissen, wer im Frühling besonders anfällig ist für den Kauf eines neuen Shampoos, das ihr Haar glänzend und für den frisch eingezogenen Junggesellen von nebenan attraktiv erscheinen lässt? Dafür waren in unserer Kultur seit den 60er Jahren die Marketing-Unternehmen zuständig. Die konnten zwar keine Daten kaufen und auswerten, sie produzierten die Daten gleich selber, mit Versuchsgruppen und Umfragen, die zuverlässige und repräsentative Ergebnisse lieferten. Man kann die Resultate so zwar nicht völlig individualisieren, und damit Einzelne gezielt bearbeiten, dafür ist die Methode aber effizienter, weil man nicht vor einem riesigen Dschungel von allgemeinen Daten steht, aus dem man nur recht mühsam die gewünschten Resultate ziehen kann.