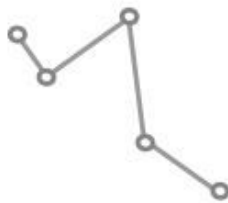


Fact Sheet

DATA LITERACY



DATA LITERACY

“Data Literacy ist die Fähigkeit, Daten zu verwenden und zu analysieren (Datenkompetenz).” - Data-Pop Alliance, 2015

Merkmale von Datenkompetenz:

- **Wissen, wie man Informationen auf verschiedene Arten findet.** Dazu gehört, in der Lage zu sein, vorhandene Datenquellen aufzuspüren.
- **Kritisches Denken** auf Daten anzuwenden. Dies reicht von der Fähigkeit zur Bewertung der Datenqualität oder zur Kontextualisierung spezifischer Informationen bis zu der Ethik im Umgang mit Daten.
- **Fragen an die Daten stellen**, auch wenn diese nicht immer die gesuchten Antworten liefern.
- **Grundlegende statistische Analysen** mit Daten durchführen können. Dies betrifft nicht nur technische Aspekte der Datenkompetenz (z. B. die Notwendigkeit Daten zu bereinigen), sondern auch die Fähigkeit, mit grundlegenden Statistiken zu arbeiten.
- **Daten zielgerichtet einsetzen** und entsprechend präsentieren, z. B. in datengesteuerten Geschichten oder Kampagnen.

Die Data Pipeline



Die Data Pipeline verschafft einen guten Überblick darüber, wie man beim Arbeiten mit Daten von der ersten Fragestellung bis zur Kommunikation vorgeht. Wichtig ist hierbei: Die Data Pipeline beschreibt **keinen linearen Vorgang**, sondern einen Prozess, in dem es nötig sein kann, einzelne Schritte zu wiederholen oder zurückzugehen und neu zu beginnen.

Die wichtigsten Schritte im Detail:



Daten finden

Im ersten Schritt werden die passenden Daten zur Fragestellung gesucht. Dazu gibt es eine Vielzahl von Tools und Techniken, angefangen von Suchmaschinen (z. B. Google-Suchoperatoren) bis hin zu Behörden, offenen Datenportalen oder einer Anfrage nach dem Informationsfreiheitsgesetz (IFG). Die richtigen Daten zu finden, kann manchmal äußerst schwierig und frustrierend sein, denn nicht alle Informationen sind frei verfügbar oder aber die Daten fehlen komplett. In diesen Fällen hilft leider nur hartnäckig zu bleiben und weiter nachzufragen, Daten selbst zu erheben (wenn möglich) oder die Fragestellung noch einmal anzupassen.



Daten säubern

Sobald alle Daten gesammelt sind, müssen sie vor der Analyse bereinigt werden. Für die Aufbereitung und alle weiteren Bearbeitungsschritte eignen sich Softwareprogramme wie OpenRefine, Excel, Libre Office oder Open Office. Die Säuberung der Daten im Vorfeld der Analyse ist wichtig, da sich insbesondere in Datensätzen, die manuell eingetragene Daten enthalten oder aus bereits vorhandenen Datensätzen neu zusammengestellt wurden, schnell kleine Fehler einschleichen.



Daten analysieren

Sobald die Daten gesäubert vorliegen, kann die Analyse starten. Auch hierfür gibt es verschiedene Tools, einfache Analyse sind aber bereits in Excel mit Hilfe von Pivot-Tabellen möglich. Mit Hilfe dieser Tabellen lassen sich Zusammenhänge in den Daten, Muster, Trends und gegebenenfalls auch Ausreißer entdecken.



Daten visualisieren

Die neu gewonnenen Informationen können nun noch grafisch aufbereitet werden. Dies ist wichtig, da gute Datenvisualisierungen im besten Fall einen leichten Zugang zur Thematik bieten und eine höhere Überzeugungskraft besitzen als Daten in einer Tabelle. Eine gute Grafik unterstreicht dabei immer die Hauptaussagen eines Textes oder einer Datenanalyse. Dazu gibt es verschiedene Tools wie z. B. infogr.am., Rawgraph.io oder Datawrapper, mit denen interaktive Datenvisualisierungen online erstellt werden können.

Quellen und weitere Ressourcen:

- Data-Pop Alliance <http://datapopalliance.org/item/what-is-data-literacy/>
- School of Data
<https://schoolofdata.org/2016/01/08/research-results-part-1-defining-data-literacy/>
- Datenschule Lernsektion
<https://datenschule.de/lernmaterialien/lernsektionen/bildungsdaten/>