

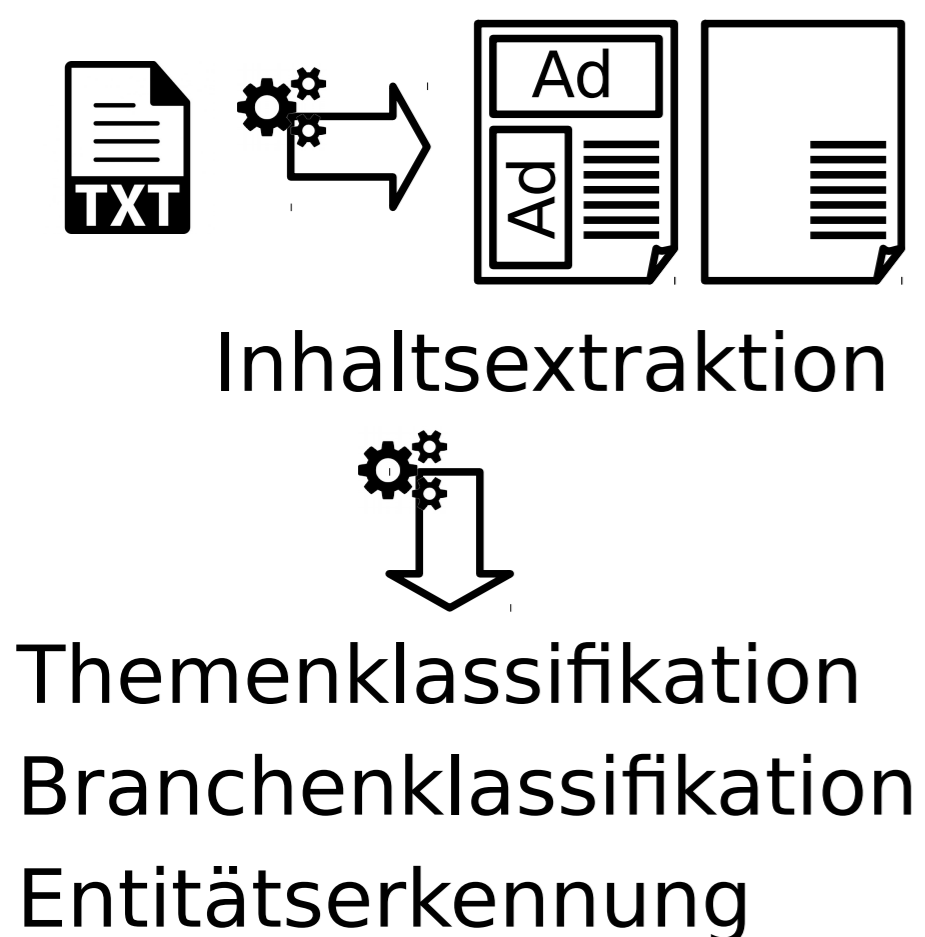
Automatisiertes vs. manuelles Data Engineering im Datenjournalismus

Abstract Auch bei Journalisten hält Data Engineering einzug. Doch gibt es keine Wunderlösung für alle. Während Reporter im Newsroom unter Zeitdruck stets über die neusten und letzten Entwicklungen berichten wollen, können Investigativjournalisten tiefer auch in heterogene Datenquellen eintauchen, sie von Grund auf analysieren und miteinander korrelieren. Zwei gänzlich unterschiedliche Ansätze vorgestellt.

Automatisierung im Newsroom



Datenanalyse



Permanent und automatisiert

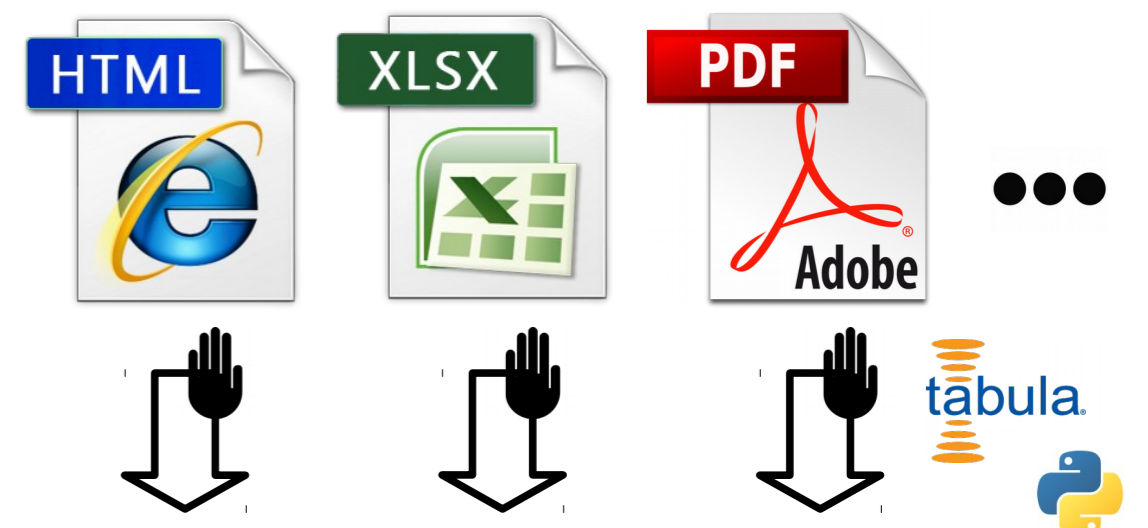
werden fremde Meldungen und Sendungen analysiert, um Redakteuren mit Übersichten zu brandaktuellen Themen und Entwicklungen, Quellen und Zitaten sowie eines Rechercheinterfaces schnelle Berichterstattung zu erleichtern.



Investigative Datenrecherche

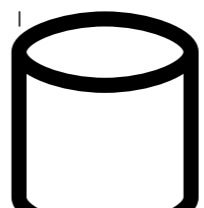
Am Beispiel von 

Open Data oder traditionelle Quellen

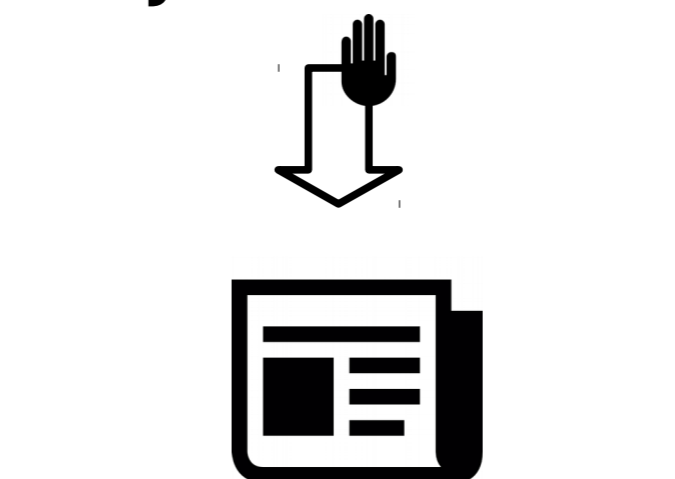


—	—	42	—	kat1	—	—	—	01010	—
—	i3e7	—	—	cg34	—	http://	—	—	—

Manuelles Data Cleansing



id1	256	berlin	peter	1327€	01001	http://	lon/lat
id2	34	köln	tim	5342€	10011	http://	lon/lat



Einmalig manuell

wird jede der verschiedenen Datenquellen im Zuge einer Recherche untersucht. Die Einrichtung von dauerhaften Pipelines findet nicht statt.

