

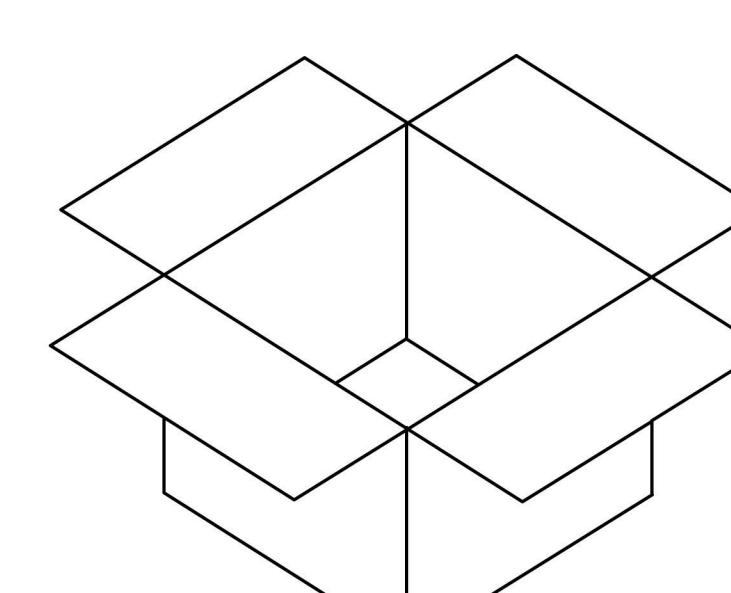
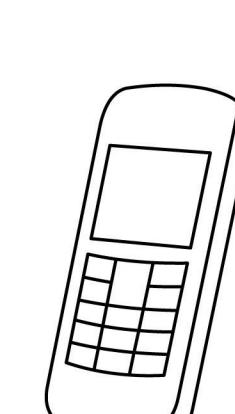
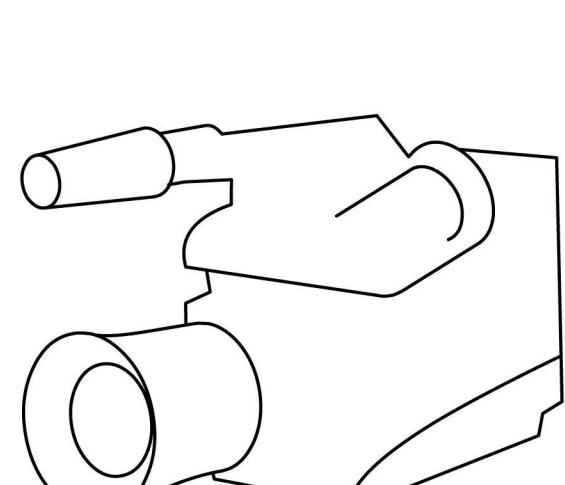
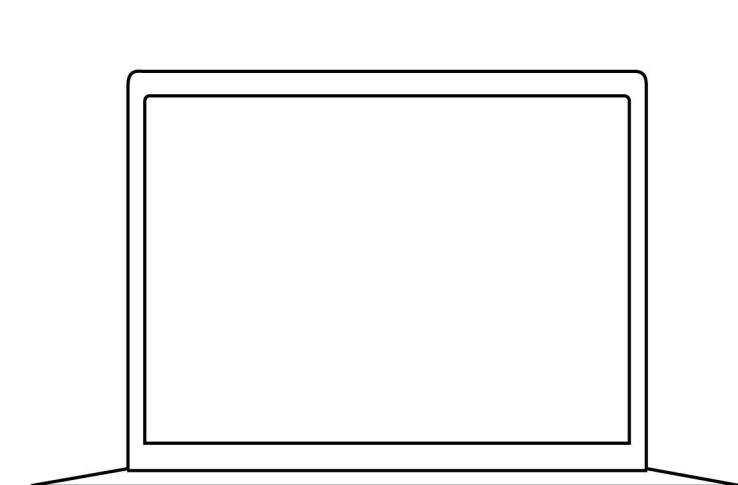
Auflösung Dokumentenaufbereitung
Kontrast Named Entity Recognition Graustufen Helligkeit
Image Processing Standardisierung
Open Source Tesseract UML API
Tokenisierung Java Tagged Image File Format
Part-of-Speech Tagging Optical Character Recognition
Stopwords ZBMED **Alzheimer-Dokumente**
FIFO-Queue Stanford Core NLP
Reguläre Ausdrücke Phrase Recognition Stemming
Word Sense Disambiguation Training von Tesseract
dynamische Textsammlung

Aufbereitung von Alzheimer-Dokumenten
für Text Mining

Entstanden im Praxisprojekt 2015

Betreut von Prof. Fischer

Ausgearbeitet von Sascha Lemke



Viel drin

Medieninformatik an der Fachhochschule Köln