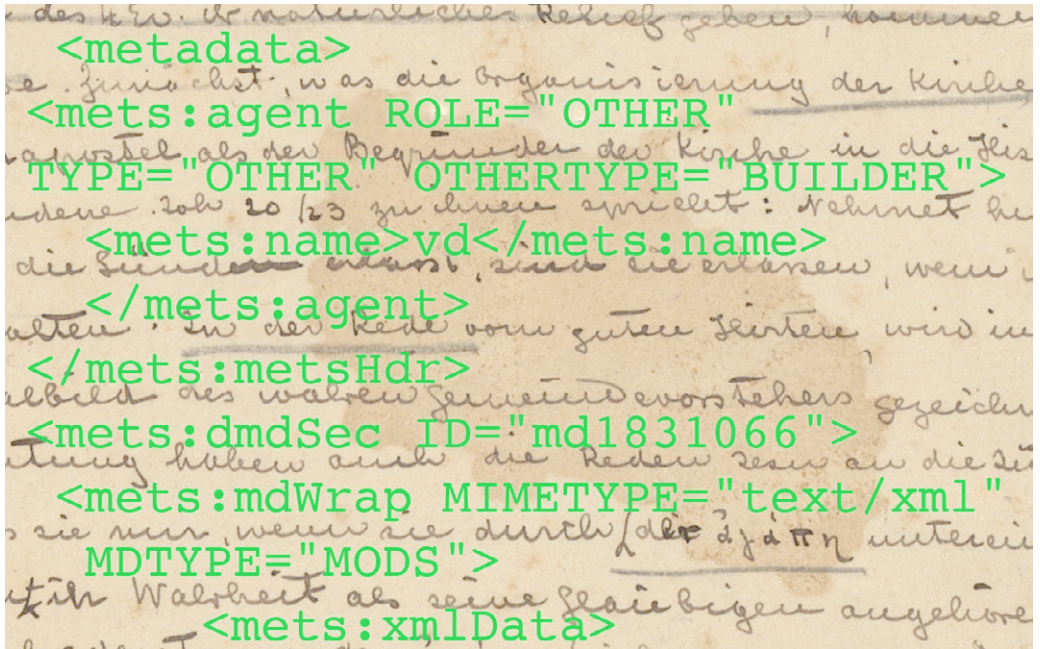


Veranstaltungsreihe «Aus der digitalen Werkstatt»



Workshop: Einführung in Theorie und Praxis der OCR mit neuronalen Netzwerken

Dozenten: Simon Clematide und Phillip Ströbel (Institut für Computerlinguistik UZH)

- Dienstag, 31. August 2021, 17.15 Uhr
- Zentralbibliothek Zürich, Hermann-Escher-Saal
- Eintritt frei

Anmeldung bis zum 24. August 2021 an: jesko.reiling@zb.uzh.ch

Einführung in Theorie und Praxis der Optical Character Recognition (OCR) mit neuronalen Netzwerken

Optical Character Recognition (OCR), die automatische Texterkennung, ist eine der wichtigsten Methoden im Bereich der Retrodigitalisierung. Dadurch werden grosse Textmengen maschinell lesbar und durchsuchbar gemacht. Insbesondere die Texterkennung von gedruckten historischen Dokumenten weist Erfolge auf, bleibt aber nach wie vor anspruchsvoll: Frakturschrift, Flecken auf den Seiten, durchscheinender Text sowie verschiedene Schriftgrößen führen dazu, dass ältere OCR-Systeme oft unverständliche und verstümmelte Texte extrahieren. Hier funktioniert die Textsuche anschliessend nur schlecht. Moderne auf neuronalen Netzwerken basierende OCR-Systeme schaffen neuerdings Abhilfe und sind in der Lage, Buchstaben zuverlässig und korrekt zu erkennen.

In unserem 90-minütigen Workshop erklären wir im ersten Teil die theoretischen und technischen Grundlagen von OCR, im zweiten Teil wenden wir uns der Praxis zu. Mit Hilfe des Tools «Transkribus» können die Teilnehmenden digitale Dokumente für die Texterkennung vorbereiten und erlernen zudem, wie das Trainieren eines eigenen Modells funktioniert.

Es wird empfohlen, für den Praxisteil der Veranstaltung den eigenen Laptop mitzubringen. Vorgängig wird darauf die notwendige Übungssoftware zu installieren sein. Eine entsprechende Anleitung wird an alle Teilnehmende rechtzeitig vor Kursbeginn verschickt.

Eine Teilnahme am Workshop ist mit vorangehender Anmeldung bis zum 24. August 2021 an diese Adresse möglich: jesko.reiling@zb.uzh.ch

In der Veranstaltungsreihe «Aus der digitalen Werkstatt. Zu Theorie und Praxis der Digitalisierung» werden zwei bis drei Mal jährlich ausgesuchte Themen aus dem Bereich der Digitalisierung erörtert und gemeinsam diskutiert.

Weitere Informationen: PD Dr. Jesko Reiling, Produktionsmanager Digitalisierung, jesko.reiling@zb.uzh.ch

Zentralbibliothek Zürich, Zähringerplatz 6, 8001 Zürich, www.zb.uzh.ch